



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 116605

TO: Jeffrey Parkin
Location: rem/3d39/3c18
Art Unit: 1648
March 25, 2004

Case Serial Number: 10/076421

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

116605

From: Parkin, Jeffrey
Sent: Wednesday, March 10, 2004 10:14 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 10/076,421

Would you please search **SEQ ID NOS.: 1-5**, from **U.S. Serial No. 10/076,421**, v. all relevant databases, including interference.

Place results on BOTH paper and electronic format.

Thanks!

JSP
AU 1648
REM 3D39
2-0908

RECEIVED
MAR 11 2004
STIC

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 3/25/04
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using SW model

(without alignments)
10505.437 Million cell updates/sec

Title: US-10-076-421-1

Perfect score: 1296
Sequence: 1 atgagagccctgcgcgcg.....agatgcccgcctctga 1296

Scoring table: IDENTITY_NUC

Gapor 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1:	gb_ba:*
2:	gb_hcg:
3:	gb_in:*
4:	gb_ov:*
5:	gb_ov:
6:	gb_pac:*
7:	gb_ph:
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sfc:
12:	gb_sy:
13:	gb_un:
14:	gb_v1:*
15:	em_ba:
16:	em_fun:
17:	em_hum:
18:	em_in:
19:	em_in:
20:	em_om:
21:	em_oz:
22:	em_ov:
23:	em_pat:
24:	em_ph:
25:	em_pl:
26:	em_ro:
27:	em_sfc:
28:	em_un:
29:	em_v1:
30:	em_hcg_hum:
31:	em_hcg_inv:
32:	em_hcg_other:
33:	em_hcg_mus:
34:	em_hcg_pln:
35:	em_hcg_rod:
36:	em_hcg_mam:
37:	em_hcg_vrt:
38:	em_sy:
39:	em_hcg_hum:
40:	em_hcg_mus:
41:	em_hcg_other:

Pred. No. is the number of results predicted by chance to have a

RESULT 1				
AX512750				
LOCUS	AX512750	1296 bp	DNA	linear
DEFINITION	Sequence 1 from Patent EP1232755..			
ACCESSION	AX512750			
VERSION	AX512750.1	GI:23503968		
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1			
AUTHORS	Wada, M. and Wada, N.			
TITLE	Anti-hiv agents			
JOURNAL	Patent: EP 1232755-A 1 21-AUG-2002;			

Db 61 AGCAATGAACTTCATCAAGTTCATCGAACTGTGACTGTCTAAATGAGAGCAATGATG 120
 Qy 121 TCCAAAGAACTTTCTCCAACTTTCACTGTGTGCACTGCCCAAGAAATTCGAGGGCAG 180
 Db 121 TCCAAAGAACTTTCTCCAACTTTCACTGTGTGCACTGCCCAAGAAATTCGAGGGCAG 180
 Qy 181 CACTGTGAATATGATATGATCAAAACCTGTGTAGAGGGGAATGATCACTTTTACGAGGA 240
 Db 181 CACTGTGAATATGATATGATCAAAACCTGTGTAGAGGGGAATGATCACTTTTACGAGGA 240
 Qy 241 AAGGCGAGACTGACACCATGAGGCGGCGCTGCTGCTGCAACTCTGCACTGTCTT 300
 Db 241 AAGGCGAGACTGACACCATGAGGCGGCGCTGCTGCTGCAACTCTGCACTGTCTT 300
 Qy 301 CAGCAACGTACATGATCCCAAGATCTGATGCTTTTCACTGTGAGGCGGGAACAATAT 360
 Db 301 CAGCAACGTACATGATCCCAAGATCTGATGCTTTTCACTGTGAGGCGGGAACAATAT 360
 Qy 361 TACTGAGGAAACCCAGACCAACCGGAGGCGACCTGTGCTATGTGCAAGTGGGCTTAAG 420
 Db 361 TACTGAGGAAACCCAGACCAACCGGAGGCGACCTGTGCTATGTGCAAGTGGGCTTAAG 420
 Qy 421 CCGCTTGTCCAGAGTGCATGTGTGATGACTGTGCGAGATGGAAGAAAGCCCTCTCTCT 480
 Db 421 CCGCTTGTCCAGAGTGCATGTGTGATGACTGTGCGAGATGGAAGAAAGCCCTCTCTCT 480
 Qy 481 CCGAAGAAATTTAAATTTTCACTGTGTGCGCAAAACCTCTGAGGCGGCTTTAAGATTAT 540
 Db 481 CCGAAGAAATTTAAATTTTCACTGTGTGCGCAAAACCTCTGAGGCGGCTTTAAGATTAT 540
 Qy 541 GGGGGAGAAATTTACCAACATCGAAGACCGGCTGTGCGGCGCATCTTACAGAGGCGAC 600
 Db 541 GGGGGAGAAATTTACCAACATCGAAGACCGGCTGTGCGGCGCATCTTACAGAGGCGAC 600
 Qy 601 CCGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGCGCTCATCAAGCCCTTGTGCGGTATC 660
 Db 601 CCGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGCGCTCATCAAGCCCTTGTGCGGTATC 660
 Qy 661 AAGGCGCAACACTGTCTTCACTTATGTTACCCAAAGAAAGAGACTTACATCTGTGCGGT 720
 Db 661 AAGGCGCAACACTGTCTTCACTTATGTTACCCAAAGAAAGAGACTTACATCTGTGCGGT 720
 Qy 721 CGCTCAAGGCTTAACCTCAACACGCAAGGGGAGATGAGTTGAGTGGAGAAACCTTAATC 780
 Db 721 CGCTCAAGGCTTAACCTCAACACGCAAGGGGAGATGAGTTGAGTGGAGAAACCTTAATC 780
 Qy 781 CTACACAAAGAGACTACAGCGCTGCAAGCTGTCTCAACCAAGACATTTGCTTGTGAGAG 840
 Db 781 CTACACAAAGAGACTACAGCGCTGCAAGCTGTCTCAACCAAGACATTTGCTTGTGAGAG 840
 Qy 841 ATCCGTTCCAAAGAGGCGAGGTGTGCGACGCAATCCCGGACTATACAGACATCTCTCTG 900
 Db 841 ATCCGTTCCAAAGAGGCGAGGTGTGCGACGCAATCCCGGACTATACAGACATCTCTCTG 900
 Qy 901 CCTCGATGTATTAAGATCCCGAGTTGGACAAGGTGTGAGATCACTGGCTTTGAGAAA 960
 Db 901 CCTCGATGTATTAAGATCCCGAGTTGGACAAGGTGTGAGATCACTGGCTTTGAGAAA 960
 Qy 961 GAGAAATTTACCGACTATCTATCCGAGACGCTGAAAATGACTTTTGAAGCTGATT 1020
 Db 961 GAGAAATTTACCGACTATCTATCCGAGACGCTGAAAATGACTTTTGAAGCTGATT 1020
 Qy 1021 TCCCAACCGGAGAGTGTGAGAGCCCACTACAGGCTGTGAGATCAACCAAAATGCTG 1080
 Db 1021 TCCCAACCGGAGAGTGTGAGAGCCCACTACAGGCTGTGAGATCAACCAAAATGCTG 1080
 Qy 1081 TGTGCTGTGACCCACAGTGAAGAAAGATTTCTGCAAGGAGACTTCAAGGGGAGCCCTTC 1140
 Db 1081 TGTGCTGTGACCCACAGTGAAGAAAGATTTCTGCAAGGAGACTTCAAGGGGAGCCCTTC 1140
 Qy 1141 GTCTGTTCCTTCCAAAGCGGCACTTGTGATGGAATTTGTGAGCTTGGGGCCGTGATGT 1200
 Db 1141 GTCTGTTCCTTCCAAAGCGGCACTTGTGATGGAATTTGTGAGCTTGGGGCCGTGATGT 1200

Qy 1201 GCCCTGAGAGACAGGCGGCGCTACAGAGAGTCTCACTTTTACCCTGATCCG 1260
 Db 1201 GCCCTGAGAGACAGGCGGCGCTACAGAGAGTCTCACTTTTACCCTGATCCG 1260
 Qy 1261 AGTCACCAAGAGAGAGAGATGCGTGGCCCTTGA 1296
 Db 1261 AGTCACCAAGAGAGAGAGATGCGTGGCCCTTGA 1296
 RESULT 3
 BD275726 2294 bp DNA linear PAT 17-JUL-2003
 LOCUS BD275726
 DEFINITION COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS OF LUNG CANCER.
 ACCESSION BD275726
 VERSION BD275726.1 GI:33085494
 KEYWORDS JP 2002543769-A/120.
 SOURCE unidentified
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 2294)
 AUTHORS Wang, T. and Van, R.
 TITLE COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS OF LUNG CANCER
 JOURNAL Patent: JP 2002543769-A 120 24-DEC-2002;
 Corixa Corporation et al
 COMMENT OS Homo sapien
 PN JP 2002543769-A/120
 PD 24-DEC-2002
 PE 03-APR-2000 JP 2000611554
 PR 22-FEB-2000 US 09/510376, 10-JAN-2000 US 09/480884, PR
 02-APR-1999 US 09/285479, 30-DEC-1999 US 09/476496, PR
 17-DEC-1999 US 09/466396
 FI longlong wang, riken van
 CC
 FEATURES
 source FH Key Location/Qualifiers.
 1..2294
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 100.0%; Score 1296; DB 6; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 1296; Conservative 0;

Qy 1 ATGAGAGCCCTCTGCGGCGGCTTCTGTGCGTCTGTGCGTGAAGCACTCAAGGC 60
 Db 67 ATGAGAGCCCTCTGCGGCGGCTTCTGTGCGTCTGTGCGTGAAGCACTCAAGGC 126
 Qy 61 AGCAATGAACTTCATCAAGTTCATCACTGTGACTGTCTTAAATGAGAAACATGTGTG 120
 Db 127 AGCAATGAACTTCATCAAGTTCATCACTGTGACTGTCTTAAATGAGAAACATGTGTG 186
 Qy 121 TCCAAACAATCTTCTCCAAACATTCATCTGTGCACTGTGCCAAAGAAATTCGAGGGCAG 180
 Db 187 TCCAAACAATCTTCTCCAAACATTCATCTGTGCACTGTGCCAAAGAAATTCGAGGGCAG 246
 Qy 181 CACTGTGAATATGATATGATCAAAACCTGTGTGAGGGGAATGATCACTTTTACGAGGA 240
 Db 247 CACTGTGAATATGATATGATCAAAACCTGTGTGAGGGGAATGATCACTTTTACGAGGA 306
 Qy 241 AAGGCGAGACTGACACCAATGAGGCGGCGCTGCTGCTGCAACTCTGCCACTGTCTT 300
 Db 307 AAGGCGAGACTGACACCAATGAGGCGGCGCTGCTGCTGCAACTCTGCCACTGTCTT 366
 Qy 301 CAGCAACGTACATGATCCCAAGATCTGATGCTTTTCACTGTGAGGCGGGAACAATAT 360
 Db 367 CAGCAACGTACATGATCCCAAGATCTGATGCTTTTCACTGTGAGGCGGGAACAATAT 426
 Qy 361 TACTGAGGAAACCCAGACCAACCGGAGGCGACCTGTGCTATGTGCAAGTGGGCGCTTAAG 420
 Db 427 TACTGAGGAAACCCAGACCAACCGGAGGCGACCTGTGCTATGTGCAAGTGGGCGCTTAAG 486

OY		421	CGGCTTGTCCAAAGATGCATGGTGCACTACTGCCAATGGAAAAAACCCCTCTCTCT	480
Db		487	CGCCTTTGTCCAAAGATGCATGGTGCACTACTGCCAATGGAAAAAACCCCTCTCTCT	546
OY		481	CCAGAAGAAATTAAAATTTCAGTGTGGCCCCAAAAGACTCGAGAGCCCCTTTAAGATTAT	540
Db		547	CCAGAAGAAATTAAAATTTCAGTGTGGCCCCAAAAGACTCGAGAGCCCCTTTAAGATTAT	606
OY		541	GGGGGAAGAAATTCACCAACCATTTGAGAACCAAGCCCTGTGTTGGCGCATCTACAGAGCAC	600
Db		607	GGGGGAAGAAATTCACCAACCATTTGAGAACCAAGCCCTGTGTTGGCGCATCTACAGAGCAC	666
OY		601	CGGGGGGGCTGTGTCAACCTAAGTGTGGAGGGACCCCTCATAGCCCTTGTCTGGTGATC	660
Db		667	CGGGGGGGCTGTGTCAACCTAAGTGTGGAGGGACCCCTCATAGCCCTTGTCTGGTGATC	726
OY		661	AGCCCCACACACTGCTCTTATTGATTATTCACCAAGAGAGAGACTACATTCGTACTGGGT	720
Db		727	AGCCCCACACACTGCTCTTATTGATTATTCACCAAGAGAGAGACTACATTCGTACTGGGT	786
OY		721	CGCTCAAGGCTTAACTCCAACGSCAAGGGGAGATGAAGTTTAAGGTGAAAACTTAATC	780
Db		787	CGCTCAAGGCTTAACTCCAACGSCAAGGGGAGATGAAGTTTAAGGTGAAAACTTAATC	846
OY		781	CTACACAAGACTACAGCGCTGAACGCTTGTCTACCAACAAGCACTTGCCTTGGCTGAAG	840
Db		847	CTACACAAGACTACAGCGCTGAACGCTTGTCTACCAACAAGCACTTGCCTTGGCTGAAG	906
OY		841	ATCCGTTTCCAAAGAGGGGAGGTGTGGGACAGCCATCCCGGACTATACAGCACTTGCCTG	900
Db		907	ATCCGTTTCCAAAGAGGGGAGGTGTGGGACAGCCATCCCGGACTATACAGCACTTGCCTG	966
OY		901	CCTCGATGTATAAGATCCCAAGTTTGGCACAAGCTGTGAGATCACTGCTTTGGAAAA	960
Db		967	CCTCGATGTATAAGATCCCAAGTTTGGCACAAGCTGTGAGATCACTGCTTTGGAAAA	1026
OY		961	GAGAATTCTAACCGACTATCTATCCGAGCGACCTGAAATGACTGTTGTGAAGCTGATT	1020
Db		1027	GAGAATTCTAACCGACTATCTATCCGAGCGACCTGAAATGACTGTTGTGAAGCTGATT	1086
OY		1021	TCCACCCGGAGTGTCAAGCAGCCCACTACTACGGCTCTGAAGTCAACCAAAATCTG	1080
Db		1087	TCCACCCGGAGTGTCAAGCAGCCCACTACTACGGCTCTGAAGTCAACCAAAATCTG	1146
OY		1081	TGTGCTGCTGACCCCAAGTGGAAAAACAATTCTGTCCACGSAAGACTCAGGGGAGCCCTC	1140
Db		1147	TGTGCTGCTGACCCCAAGTGGAAAAACAATTCTGTCCACGSAAGACTCAGGGGAGCCCTC	1206
OY		1141	GTTGTGTTCCCTTCAAAGCCGCGATGACTTTGACTGGAATTGTGAAGCTGGGGCCGTGATGT	1200
Db		1207	GTTGTGTTCCCTTCAAAGCCGCGATGACTTTGACTGGAATTGTGAAGCTGGGGCCGTGATGT	1266
OY		1201	GGCCGGAAGGACAAAGCCAGCGCTCTACACGAGAGTTTACACTTTACCCTGGATCCGC	1260
Db		1267	GGCCGGAAGGACAAAGCCAGCGCTCTACACGAGAGTTTACACTTTACCCTGGATCCGC	1326
OY		1261	AGTCACACCAAGAAAGAAATGAGCTGGCCCTCTGA	1296
Db		1327	AGTCACACCAAGAAAGAAATGAGCTGGCCCTCTGA	1362
RESULT 4				
LOCUS	AR220511			
DEFINITION	Sequence 123 from patent US 6426072.	2294 bp	DNA	linear
ACCESSION	AR220511			PAT 26-SEP-2002
VERSION	AR220511.1			
KEYWORDS	GI:23327292			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
	1 (bases 1 to 2294)			

AUTHORS	Wang, T. T., Fan, L., Katos, M. D., Wang, C. S., Hosken, N. A., Fanger, G. R., Li, S. X., Wang, A., Seikely, Y. A. M., Henderson, R. A. and McNeill, P. D.
TITLE	Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL FEATURES	Patent: US 6426072-A 123 30-UU-2002;
source	Location/Qualifiers 1. .2294
ORIGIN	/Organization="unknown" /mol_type="genomic DNA"

Query Match	100.0%;	Score 1296;	DB 6;	Length 2294;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1296;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	ATGAGAGCCCTGTGGGCGGCGGGTTCTCTGGCTCTTGGTCTGTGACGACTCCAAAGC	6
Db	67	ATGAGAGCCCTGTGGGCGGCGGGTTCTCTGGCTCTTGGTCTGTGACGACTCCAAAGC	12
Qy	61	AGCAATGAACCTTCATCAAGTTCCATCGAACCTGTACTGTCTAAATGGAGGACATATGTC	12
Db	127	AGCAATGAACCTTCATCAAGTTCCATCGAACCTGTACTGTCTAAATGGAGGACATATGTC	18
Qy	121	TCCAAACAAGTACTTCTCCAAACATTCACCTGGTGCACCTGCCAAAGAAATTCGAGGGCAG	18
Db	187	TCCAAACAAGTACTTCTCCAAACATTCACCTGGTGCACCTGCCAAAGAAATTCGAGGGCAG	24
Qy	181	CACCTGTGAATTAAGTAAGTCAAAAACTGTCTATGAGGGGAAATGTCACTTTTACCGAGA	24
Db	247	CACCTGTGAATTAAGTAAGTCAAAAACTGTCTATGAGGGGAAATGTCACTTTTACCGAGA	30
Qy	241	AAGGCACACACTGCACACATGAGCGCGGCGCTGTGCTCCCTGGAACCTGTGCCACTGCTT	30
Db	307	AAGGCACACACTGCACACATGAGCGCGGCGCTGTGCTCCCTGGAACCTGTGCCACTGCTT	36
Qy	301	CAGCAAACTTAACATATGCCCAAGATCTGATGCTTTCAAGCTGAGCGCTGGGGGAAACATAT	36
Db	367	CAGCAAACTTAACATATGCCCAAGATCTGATGCTTTCAAGCTGAGCGCTGGGGGAAACATAT	42
Qy	361	TACTGCAGGAAACCCAGACACCGGAGGCGACCCCTGTGTATGTGCAAGTGGGCTTAAG	42
Db	427	TACTGCAGGAAACCCAGACACCGGAGGCGACCCCTGTGTATGTGCAAGTGGGCTTAAG	48
Qy	421	CCGCTTGTCCAGAGTGCATGTGCATGACTGCGAGATGGAAGAAACCCCTCTCTCT	48
Db	487	CCGCTTGTCCAGAGTGCATGTGCATGACTGCGAGATGGAAGAAACCCCTCTCTCT	54
Qy	481	CCAGAAATTAATAATTTTCAGTGTGGCCAAAAGACTCTGAGGCCCGCCCTTTAAGATTAT	54
Db	547	CCAGAAATTAATAATTTTCAGTGTGGCCAAAAGACTCTGAGGCCCGCCCTTTAAGATTAT	60
Qy	541	GGGGAGAAATTCACCAACATGAGAACCAAGCCCTGTGTGGGCCATCTACAGAGGAC	60
Db	607	GGGGAGAAATTCACCAACATGAGAACCAAGCCCTGTGTGTGGGCCATCTACAGAGGAC	66
Qy	601	CGGGGGGGGCTCTGTACCTTAAGTGTGTGAGGAGAGCTTCATCAGCCCTTGTGTGTATC	66
Db	667	CGGGGGGGGCTCTGTACCTTAAGTGTGTGAGGAGAGCTTCATCAGCCCTTGTGTGTATC	72
Qy	661	AGGCCACACACTGCTTCAATGATTATCCCAAAGAGAGACTACATCGTCTACCTGGGT	72
Db	727	AGGCCACACACTGCTTCAATGATTATCCCAAAGAGAGACTACATCGTCTACCTGGGT	78
Qy	721	CGCTCAAGGCTTAATCCCAACCGCAAGGGGAGATGAAGTTTGAAGTGGAAAACTTAATC	78
Db	787	CGCTCAAGGCTTAATCCCAACCGCAAGGGGAGATGAAGTTTGAAGTGGAAAACTTAATC	84
Qy	781	CTACACAAAGACTACAGGCTGACACAGCTTGCTCACCAACAAGCAATTTGCTTGCTGAAG	84
Db	847	CTACACAAAGACTACAGGCTGACACAGCTTGCTCACCAACAAGCAATTTGCTTGCTGAAG	90
Qy	841	ATTCCTTCCAAAGAGGCGAGGTGTGGCAGCAATCCCGGACTATACAGACATATGCTGTG	90

Db 907 ATCCGTTCCAAAGGAGGAGAGTGTGCGAGCCATCCCGAGCTATPACAGACCATCTGCTG 966
Qy 901 CCTCGATGATTAACATATCCCGAGTTTGGACAAGCTGTAGATCACTGGCTTTGGAAAA 960
Db 967 CCTCATGATATACATATCCCGAGTTTGGACAAGCTGTAGATCACTGGCTTTGGAAAA 1026
Qy 961 GAGAAATTTACCGACTATCTATCCGAGCAGCTGAAAAATGACTGTGTGGAAGCTGATT 1020
Db 1027 GAGAAATTTACCGACTATCTATCCGAGCAGCTGAAAAATGACTGTGTGGAAGCTGATT 1086
Qy 1021 TCCCAACGGGAGTGTGAGCAGCCCACTACTACGGCTGTGAAGTCAACAACAAATGCTG 1080
Db 1087 TCCCAACGGGAGTGTGAGCAGCCCACTACTACGGCTGTGAAGTCAACAACAAATGCTG 1146
Qy 1081 TGTGCTGCTGACCCCAACAGTGTGAAAAAGATTTCTGCCAGGAGACTCAGGGGGACCCCTC 1140
Db 1147 TGTGCTGCTGACCCCAACAGTGTGAAAAAGATTTCTGCCAGGAGACTCAGGGGGACCCCTC 1206
Qy 1141 GTCTGTTCCCTCCCAAGCCGCGATGACTTTGACTGSAATTGTGAGCTGGGGCCGTTGATGT 1200
Db 1207 GTCTGTTCCCTCCCAAGCCGCGATGACTTTGACTGSAATTGTGAGCTGGGGCCGTTGATGT 1266
Qy 1201 GCCCTGAAGAGACAAAGCCAGGCGCTCTACACAGAGTCTCAACCTTCTTACCCTGATCCGC 1260
Db 1267 GCCCTGAAGAGACAAAGCCAGGCGCTCTACACAGAGTCTCAACCTTCTTACCCTGATCCGC 1326
Qy 1261 AGTCACACCAAGAGAGAAATGGCCCTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAAATGGCCCTGGCCCTCTGA 1362

RESULT 5
AR255505
LOCUS AR255505 2294 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 123 from patent US 6482597.
ACCESSION AR255505
VERSION AR255505.1 GI:27304554
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2294)
AUTHORS Wang, T., Hosken, N.A., Kalos, M.D., Fanger, G.R. and Fan, L.
TITLE Compounds and methods for therapy and diagnosis of lung cancer
JOURNAL Patent: US 6482597-A 123 19-NOV-2002;
FEATURES
source 1..2294
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGAGCCCTGTGCGCGGCGCTGCTTCTCTGCGTCTTGCTGTGCACTTCCAAAAGC 60
Db 67 ATGAGAGCCCTGTGCGCGGCGCTGCTTCTCTGCGTCTTGCTGTGCACTTCCAAAAGC 126
Qy 61 AGCAATGAAGCTTCATCAAGTTCATGAACTGTGAACTGTCTAAATGAGAAACATGTGTG 120
Db 127 AGCAATGAAGCTTCATCAAGTTCATGAACTGTGAACTGTCTAAATGAGAAACATGTGTG 186
Qy 121 TCCCAACAGTACTTCTTCCAAACATTCACCTGTGTGCACTGCCCAGAAAGAAATTCGAGGGCAG 180
Db 187 TCCCAACAGTACTTCTTCCAAACATTCACCTGTGTGCACTGCCCAGAAAGAAATTCGAGGGCAG 246
Qy 181 CACTGTGAATATGATATGTCAAAAACCTGTATGAGGGGAATGTCACTTTTACCGAGGA 240
Db 247 CACTGTGAATATGATATGTCAAAAACCTGTATGAGGGGAATGTCACTTTTACCGAGGA 306
Qy 241 AAGGCGAGACTGACCAATGAGGCGGCGCTGCTGCGCTGGAACCTGCGCACTGTCCTT 300

Db 307 AAGGCGAGACTGACACACATGAGGCGGCGCCCTGCTGCGCTTGAACTCTGCACTGTCCTT 366
Qy 301 CAGCAAAAGTATCCATATGCTCCACAGATGTGATGCTTCTTCACTGTGGGCTGTGGGAAAACTAAT 360
Db 367 CAGCAAAAGTATCCATATGCTCCACAGATGTGATGCTTCTTCACTGTGGGCTGTGGGAAAACTAAT 426
Qy 361 TACTGAGGAAACCCAGACAAACCGGAGGCGACCCCTGGTGTATGTGAGAGTGGGCTTAAG 420
Db 427 TACTGAGGAAACCCAGACAAACCGGAGGCGACCCCTGGTGTATGTGAGAGTGGGCTTAAG 486
Qy 421 CCGCTTGTCCAAAGTGCATGTGTGATGACTGTGCGAGATGAAAAAGCCCTCTCTCT 480
Db 487 CCGCTTGTCCAAAGTGCATGTGTGATGACTGTGCGAGATGAAAAAGCCCTCTCTCTCT 546
Qy 481 CCAAGAAATTAATTAATTAATGATGTGCGCAAAAGACTGTGAGGCCCGCTTTAAGTTATT 540
Db 547 CCAAGAAATTAATTAATTAATGATGTGCGCAAAAGACTGTGAGGCCCGCTTTAAGTTATT 606
Qy 541 GGGGGGAAATTCACCAACATCGAAGAACAGCCCTGTGTTGGCGGCACTTACAGAGGCAAC 600
Db 607 GGGGGGAAATTCACCAACATCGAAGAACAGCCCTGTGTTGGCGGCACTTACAGAGGCAAC 666
Qy 601 CCGGGGGGCTCTGTCACTACGTGTGTGAGAGGAGGCTCATAGCCCTTGTGTGGTATC 660
Db 667 CCGGGGGGCTCTGTCACTACGTGTGTGAGAGGAGGCTCATAGCCCTTGTGTGGTATC 726
Qy 661 AGGCGCAACACTGCTTCAATGATTAACCAAGAGAGAGACTAATCGTCTACCTGGGT 720
Db 727 AGGCGCAACACTGCTTCAATGATTAACCAAGAGAGAGACTAATCGTCTACCTGGGT 786
Qy 721 CGCTCAAGGCTTAACCTCCAAACGCAAGGGGAGATGAATTTGAAGTGAAGAAACCTAATC 780
Db 787 CGCTCAAGGCTTAACCTCCAAACGCAAGGGGAGATGAATTTGAAGTGAAGAAACCTAATC 846
Qy 781 CTACAAAGAACTACAGCGCTGACAGCTTGTCTCACACAGCACTTGGCTTGTGAAG 840
Db 847 CTACAAAGAACTACAGCGCTGACAGCTTGTCTCACACAGCACTTGGCTTGTGAAG 906
Qy 841 ATCCGTTCCAAAGAGGAGAGTGTGTGCGAGCCATCCCGGACTATACAGACCATCTGCTG 900
Db 907 ATCCGTTCCAAAGAGGAGAGTGTGTGCGAGCCATCCCGGACTATACAGACCATCTGCTG 966
Qy 901 CCTCGATGATTAACATATCCCGAGTTTGGACAAGCTGTGAGATCACTGGCTTTGGAAAA 960
Db 967 CCTCGATGATTAACATATCCCGAGTTTGGACAAGCTGTGAGATCACTGGCTTTGGAAAA 1026
Qy 961 GAGAAATTTACCGACTATCTATCCGAGCAGCTGAAAAATGACTGTGTGGAAGCTGATT 1020
Db 1027 GAGAAATTTACCGACTATCTATCCGAGCAGCTGAAAAATGACTGTGTGGAAGCTGATT 1086
Qy 1021 TCCCAACGGGAGTGTGAGCAGCCCACTACTACGGCTGTGAAGTCAACAACAAATGCTG 1080
Db 1087 TCCCAACGGGAGTGTGAGCAGCCCACTACTACGGCTGTGAAGTCAACAACAAATGCTG 1146
Qy 1081 TGTGCTGCTGACCCCAACAGTGTGAAAAAGATTTCTGCCAGGAGACTCAGGGGGACCCCTC 1140
Db 1147 TGTGCTGCTGACCCCAACAGTGTGAAAAAGATTTCTGCCAGGAGACTCAGGGGGACCCCTC 1206
Qy 1141 GTCTGTTCCCTCCCAAGCCGCGATGACTTTGACTGSAATTGTGAGCTGGGGCCGTTGATGT 1200
Db 1207 GTCTGTTCCCTCCCAAGCCGCGATGACTTTGACTGSAATTGTGAGCTGGGGCCGTTGATGT 1266
Qy 1201 GCCCTGAAGAGACAAAGCCAGGCGCTTACACAGAGTCTCAACCTTCTTACCCTGATCCGC 1260
Db 1267 GCCCTGAAGAGACAAAGCCAGGCGCTTACACAGAGTCTCAACCTTCTTACCCTGATCCGC 1326
Qy 1261 AGTCACACCAAGAGAGAAATGGCCCTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAAATGGCCCTGGCCCTCTGA 1362

RESULT 6
AR281075

187 TCACAAAGTACTTCTCCAACTTCACTGTGCACTGCCCAAGAAATTGGAGGGCAG 246
181 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAAATGGTCACTTTTACCGAGA 240
247 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAAATGGTCACTTTTACCGAGA 306
241 AAGGCGAGACTGACACACATGGGGCCGCTGCTGCTGCAATCTGTGCACTGTCTT 300
307 AAGGCGAGACTGACACACATGGGGCCGCTGCTGCTGCAATCTGTGCACTGTCTT 366
301 CAGCAAAAGTACCATGCGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAT 360
367 CAGCAAAAGTACCATGCGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAT 426
361 TACTGACAGAACCCAGCAACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
427 TACTGACAGAACCCAGCAACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 486
421 CGGCTTGTCCAAAGTGCATGTGCACTGATGCTGGGCAATGGAAAAAGCCCTCTCTCT 480
487 CGGCTTGTCCAAAGTGCATGTGCACTGATGCTGGGCAATGGAAAAAGCCCTCTCTCT 546
481 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
547 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 606
541 GGGGGAAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGCCCATCTACAGAGGAC 600
607 GGGGGAAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGCCCATCTACAGAGGAC 666
601 GGGGGAAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGCCCATCTACAGAGGAC 660
667 GGGGGAAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGCCCATCTACAGAGGAC 726
661 AAGGCGAGACTGACACACATGGGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
727 AAGGCGAGACTGACACACATGGGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
721 CGGCTAAGGCTTAATCTCCAAACAGCAACGGGAGATGAAGTTGAGGTGAAAACTTAATC 780
787 CGGCTAAGGCTTAATCTCCAAACAGCAACGGGAGATGAAGTTGAGGTGAAAACTTAATC 846
781 CTAACAAGAACTACAGCGCTGACAGCTTGTCTACCAACAGCATTTGCTTGTCTGAG 840
847 CTAACAAGAACTACAGCGCTGACAGCTTGTCTACCAACAGCATTTGCTTGTCTGAG 906
841 ATCCGTTCCAAAGGAGGAGGAGTGTGCGAGCCATCCCGGACTATACAGACATCTGCTG 900
907 ATCCGTTCCAAAGGAGGAGGAGTGTGCGAGCCATCCCGGACTATACAGACATCTGCTG 966
901 CCCTGATGTATACGATCCCAAGTTGGCAAGAGTGTGAGATCACTGGCTTTGAAAA 960
967 CCCTGATGTATACGATCCCAAGTTGGCAAGAGTGTGAGATCACTGGCTTTGAAAA 1026
961 GAGAAATTTCAACCACTATCTATCCGAGCACTGAAAAATGACTGTGTGAAAGCTGAT 1020
1027 GAGAAATTTCAACCACTATCTATCCGAGCACTGAAAAATGACTGTGTGAAAGCTGAT 1086
1021 TCCCAACGGGAGAGTGTGAGGAGCCCACTACTAGGGCTGAAATGCAACCAAAAGCTG 1080
1087 TCCCAACGGGAGAGTGTGAGGAGCCCACTACTAGGGCTGAAATGCAACCAAAAGCTG 1146
1081 TGTGTGTGTGACCCACAGTGAAGAAAGATTTCTGTCCAGGAGACTCAGGGGGAGCCCTC 1140
1147 TGTGTGTGTGACCCACAGTGAAGAAAGATTTCTGTCCAGGAGACTCAGGGGGAGCCCTC 1206
1141 GTCTGTCTCTCCAAAGGCGGAGTGTGAACTGTGAATTTGTAGCTGGGGCTGTGATGT 1200
1207 GTCTGTCTCTCCAAAGGCGGAGTGTGAACTGTGAATTTGTAGCTGGGGCTGTGATGT 1266
1201 GGCCTAAAGAACAGGCGGCTGTACAGAGAGTCTACACTTTTACCCTGATCCGCG 1260

Db 1267 GCCCTGAGAGCAACGAGGCGTCTACAGAGAGTCTACACTTTTACCCTGATCCGCG 1326
Qy 1261 AGTCACACCAAGAGAGAAATGGCTTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAAATGGCTTGGCCCTCTGA 1362
RESULT 8
AR437866
LOCUS AR437866 2294 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 123 from patent US 6660838.
ACCESSION AR437866
VERSION AR437866.1 GI:40203082
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2294)
AUTHORS Wang, T.
TITLES Compounds and methods for therapy and diagnosis of lung cancer
JOURNAL Patent: US 6660838-A 123 09-DEC-2003;
FEATURES Location/Qualifiers
source 1..2294
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAGCCCTGCTGCGCGGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 67 ATGAGAGCCCTGCTGCGCGGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
Qy 61 AGCAATGAACTTATCATGAAGTTCCATGAACTGTAAGTGTAAATGAGAGACATGTGTG 120
Db 127 AGCAATGAACTTATCATGAAGTTCCATGAACTGTAAGTGTAAATGAGAGACATGTGTG 186
Qy 121 TCCAAAGTACTTCTCCAAACATTCACGTGTGCACTGCCCAAGAAATTTGGAGGGCAG 180
Db 187 TCCAAAGTACTTCTCCAAACATTCACGTGTGCACTGCCCAAGAAATTTGGAGGGCAG 246
Qy 181 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAAATGGTCACTTTTACCGAGA 240
Db 247 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAAATGGTCACTTTTACCGAGA 306
Qy 241 AAGGCGAGACTGACACACATGGGGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 307 AAGGCGAGACTGACACACATGGGGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
Qy 301 CAGCAAAAGTACCATGCGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAT 360
Db 367 CAGCAAAAGTACCATGCGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAT 426
Qy 361 TACTGACAGAACCCAGCAACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
Db 427 TACTGACAGAACCCAGCAACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 486
Qy 421 CGGCTTGTCCAAAGTGCATGTGCACTGATGCTGGGCAATGGAAAAAGCCCTCTCTCT 480
Db 487 CGGCTTGTCCAAAGTGCATGTGCACTGATGCTGGGCAATGGAAAAAGCCCTCTCTCT 546
Qy 481 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
Db 547 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 606
Qy 541 GGGGGAAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGGCACTTACAGAGGAC 600
Db 607 GGGGGAAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGGCACTTACAGAGGAC 666
Qy 601 GGGGGAAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGGCACTTACAGAGGAC 660

Db	667	CGGGGGGGGCTCTGTCACTACGTGTGTGGAGGACGCTCATACGCCCTTGCTGGGTGATC	726
Qy	661	AGGCCACACACTGCTTCATTGATTAACCAAGAGAGAGACTACATCGTCTACCTGGGT	720
Db	727	AGGGCCACACACTGCTTCATTGATTAACCAAGAGAGAGACTACATCGTCTACCTGGGT	786
Qy	721	CGCTCAAGGCTTAACCTCAACACGCAAGGGAGATGAAGTTGAAGGTGGAAAACCTAAAC	780
Db	787	CGCTCAAGGCTTAACCTCAACACGCAAGGGAGATGAAGTTGAAGGTGGAAAACCTAAAC	846
Qy	781	CTACACAAGGACTYACAGCGCTGACACGCTTGTCTACACACACGACTTCCTTGTGTAAG	840
Db	847	CTACACAAGGACTYACAGCGCTGACACGCTTGTCTACACACACGACTTCCTTGTGTAAG	906
Qy	841	ATCCGTTCCAAAGGAGGGAGGTGTGTGGAGCCATCCCAGACTTAAAGACATCTGCTTG	900
Db	907	ATCCGTTCCAAAGGAGGGAGGTGTGTGGAGCCATCCCAGACTTAAAGACATCTGCTTG	966
Qy	901	CCCTCGATGATTAACGATCCCAAGTTGGACACAGCTGAGATCACTGGCTTTGGAAAA	960
Db	967	CCCTCGATGATTAACGATCCCAAGTTGGACACAGCTGAGATCACTGGCTTTGGAAAA	1022
Qy	961	GAGAAATTTCAACGACTATCTTAATCCGGAGCAGCTGAATAATGACTGTTGTGAAGCTAAT	1020
Db	1027	GAGAAATTTCAACGACTATCTTAATCCGGAGCAGCTGAATAATGACTGTTGTGAAGCTAAT	1086
Qy	1021	TCCACACGGGAGTGTGACGAGCCCACTACTAGCGCTCGAAGTCAACACAAATGCTG	1080
Db	1087	TCCACACGGGAGTGTGACGAGCCCACTACTAGCGCTCGAAGTCAACACAAATGCTG	1146
Qy	1081	TGTGTGCTGACCCACAGTGTGAAAAACAGATTCTCTGACAGGAGACTCAGGGGAGCCCTC	1140
Db	1147	TGTGTGCTGACCCACAGTGTGAAAAACAGATTCTCTGACAGGAGACTCAGGGGAGCCCTC	1206
Qy	1141	GTCTGTTCCCTCAAGAGCCGCAATGACTTGAATGACGTGGAGCGGGCCGTGGAGAT	1200
Db	1207	GTCTGTTCCCTCAAGAGCCGCAATGACTTGAATGACGTGGAGCGGGCCGTGGAGAT	1266
Qy	1201	GGCTGGAAGGACAAGCCAGCGGTCTACAGAGAGTCTCAACTTCTTAACCTTGATCCGC	1260
Db	1267	GGCTGGAAGGACAAGCCAGCGGTCTACAGAGAGTCTCAACTTCTTAACCTTGATCCGC	1332
Qy	1261	AGTCACACCAAGAGAGAAATGGCCCTGGCCCTCTGA	1296
Db	1327	AGTCACACCAAGAGAGAAATGGCCCTGGCCCTCTGA	1362
RESULT 9			
AX365730		2294 bp	DNA
LOCUS			
DEFINITION	Sequence 123 from Patent WO0200174.		
ACCESSION	AX365730		
VERSION	AX365730.1		
KEYWORDS			
ORGANISM			
SOURCE			
ORGANISM	Homio sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	Wang, T., Wang, A., Skeikh, Y. A., Li, S. X., Kalos, M. D., Henderson, R. A., McNeill, P. D., Fanger, N., Retter, M. W., Marnerakis, M., Fanger, G. R., Vedelick, T. S., Carter, D., Matanabe, Y. and Peckham, D. W.		
JOURNAL	Compositions and methods for the therapy and diagnosis of lung cancer		
FEATURES	Patent: WO 0200174-A 123 03-JAN-2002;		
source	CORIXA CORPORATION (US)		
ORIGIN	Location/Qualifiers		
	1..2294		
	/organism="Homio sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		

Query Match	100.0%	Score 1296	DB 6	Length 2294	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 1296	Conservative 0	Mismatch 0	Indels 0	Gaps 0	
QY	1	ATGAGAGCCCTGGCGGCGCTTCTCTGCGCTGCTGAGCGACCTCCAAAGC	60		
Db	67	ATGAGAGCCCTGGCGGCGCTTCTCTGCGCTGCTGAGCGACCTCCAAAGC	126		
QY	61	AGCAATGAATTATCAATGTTCCATGCACTGTGACTGTAAATGAGGAACATGTG	120		
Db	127	AGCAATGAATTATCAATGTTCCATGCACTGTGACTGTAAATGAGGAACATGTG	186		
QY	121	TCCAAACAATACTTCTCCCAACTTCACTGTGTCACTGTCCCAAGAAATTCGAGGCGAG	180		
Db	187	TCCAAACAATACTTCTCCCAACTTCACTGTGTCACTGTCCCAAGAAATTCGAGGCGAG	246		
QY	181	CACTGTAAATAGATTAAGTCAAAAACTGTCTATGAGGGGAATGTGTCATTTAACGAGGA	240		
Db	247	CACTGTAAATAGATTAAGTCAAAAACTGTCTATGAGGGGAATGTGTCATTTAACGAGGA	306		
QY	241	AAGGCCAGCACTGACACCATGGGCGGCGCTGTGCTGCTGTGAACTTGTCCACTGTCTT	300		
Db	307	AAGGCCAGCACTGACACCATGGGCGGCGCTGTGCTGCTGTGAACTTGTCCACTGTCTT	366		
QY	301	CAGCAAAAGTATCCATGGCCACAGATCTGATGTCTTCAAGTGGGCTGTGGGAAAAATATAT	360		
Db	367	CAGCAAAAGTATCCATGGCCACAGATCTGATGTCTTCAAGTGGGCTGTGGGAAAAATATAT	426		
QY	361	TACTGAGGAACCCAGACACCGGAGGGGACCTGTGTCTATGTGACAGTGTGGGCTTAAG	420		
Db	427	TACTGAGGAACCCAGACACCGGAGGGGACCTGTGTCTATGTGACAGTGTGGGCTTAAG	486		
QY	421	CGCTTGTCCAAAGATGTGATGTGATGATCTGCGAGATGAAAAAABCCCTCTCTCT	480		
Db	487	CGCTTGTCCAAAGATGTGATGTGATGATCTGCGAGATGAAAAAABCCCTCTCTCT	546		
QY	481	CCAGAAATTAATAATTTCACTGTGTGCGAAAAAGACTCTGAGGCGCCGCTTTAAGATTAT	540		
Db	547	CCAGAAATTAATAATTTCACTGTGTGCGAAAAAGACTCTGAGGCGCCGCTTTAAGATTAT	606		
QY	541	GGGGGGAATTCACCAACATGAGAAACCAAGCCCTGTGTGTGGGCACTTACAGAGGCAC	600		
Db	607	GGGGGGAATTCACCAACATGAGAAACCAAGCCCTGTGTGTGGGCACTTACAGAGGCAC	666		
QY	601	CGGGGGGCTGTGCACTTACGTGTGTGAGGAGCACTCATCAGCCCTGTGTGGGTATC	660		
Db	667	CGGGGGGCTGTGCACTTACGTGTGTGAGGAGCACTCATCAGCCCTGTGTGGGTATC	726		
QY	661	AGCGCAACACTGTTTATTGATTTACCAAGAAAGAGACTACATGTCTTACTGTGGGT	720		
Db	727	AGCGCAACACTGTTTATTGATTTACCAAGAAAGAGACTACATGTCTTACTGTGGGT	786		
QY	721	CGCTCAAGGCTTAACTCCCAACGCAAGGGAGATGAAGTTGAGGTGAAAAACCTATATC	780		
Db	787	CGCTCAAGGCTTAACTCCCAACGCAAGGGAGATGAAGTTGAGGTGAAAAACCTATATC	846		
QY	781	CTACACAAAGACTACAGGCTGTGACACGTTGTCTACCAACAACGACATTTGCTTGTGAG	840		
Db	847	CTACACAAAGACTACAGGCTGTGACACGTTGTCTACCAACAACGACATTTGCTTGTGAG	906		
QY	841	ATCGGTTCCAAAGGAGGGAGGTGTGGCAGCACTCCGGACTATACAGACCATTTGGCTG	900		
Db	907	ATCGGTTCCAAAGGAGGGAGGTGTGGCAGCACTCCGGACTATACAGACCATTTGGCTG	966		
QY	901	CCCTGATGTATTAACGATCCCAAGTTTGTGCAAGAGCTGTGAGATCATGTGCTTTGAAAA	960		
Db	967	CCCTGATGTATTAACGATCCCAAGTTTGTGCAAGAGCTGTGAGATCATGTGCTTTGAAAA	1026		
QY	961	GAGAAATTCACCGACTATCTTATTCGGAGAGAGCTGAAAAATGACTGTTGTGAACTGATT	1020		
Db	1027	GAGAAATTCACCGACTATCTTATTCGGAGAGAGCTGAAAAATGACTGTTGTGAACTGATT	1086		
QY	1021	TCCACCGGGAGTGTGACAGGCCCACTACTACGGCTGTGAAGTCAACCAAAATGCTG	1080		

Accession	Sequence	Position
Db	1087 TCCACCGGAGCTGTACGACGCCCCACTACTACGGCTCTGAAGTCCACCAACCAATGCTG	1146
Qy	1081 TGTGCTGTGACCCACAGTGGAAAAACAGATTCTCGCCAGGAGACTCAGGGGGACCCCTC	1140
Db	1147 TGTCTGTGTGACCCACAGTGGAAAAACAGATTCTCGCCAGGAGACTCAGGGGGACCCCTC	1206
Qy	1141 GTCTGTTCCCTCCAAAGGCCGATGACTTTGACTGGAAATTGAGCTGGGGCCGTGATGT	1200
Db	1207 GTCTGTTCCCTCCAAAGGCCGATGACTTTGACTGGAAATTGAGCTGGGGCCGTGATGT	1266
Qy	1201 GCCCTGAAGGCAACGCCAGGCGTGTACACGAGAGTCCACACTTCTTAACCTGGATCCG	1260
Db	1267 GCCCTGAAGGCAACGCCAGGCGTGTACACGAGAGTCTACACACTTCTTAACCTGGATCCG	1326
Qy	1261 AGTCAACACCAAGAAAGATAGCGCTGGCCCTCTTGA	1296
Db	1327 AGTCAACACCAAGAAAGATAGCGCTGGCCCTCTTGA	1362

RESULT 10	BD226055	2394 bp	DNA linear	PAT 17-JUL-2003
LOCUS	BD226055			
DEFINITION	Compound and method for remedy and diagnosis of lung cancer.			
ACCESSION	BD226055			
VERSION	BD226055.1	GI:33035825		
KEYWORDS	JP 2002533056-A/120.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			

REFERENCE
1 (bases 1 to 2294)
AUTHORS
Reed, S.G. and Wang, T.
TITLE
Compound and method for remedy and diagnosis of lung cancer
JOURNAL
Patent: JP 2002533056-A 120 08-OCT-2002;

COMMENT	OS	Homo sapiens (human)
	PN	JP 2002533056-A/120
	PD	08-OCT-2002
	PF	17-MAR-1999 JP 2000536857
	PR	18-MAR-1998 US 09/040802, 18-MAR-1998 US 09/040984 PR
	27-JUL-1998 US	09/123912, 27-JUL-1998 US 09/123933 PI STEVENSON
	CC	REED, TONGTONG WANG
	PC	C12N15/09, A61K31/711, A61K35/14, A61K38/00, A61K38/16, A61K39/00, A61K39/39,
	PC	A61K48/00, A61P35/00, C07K14/47, C07K16/18, C12N1/21, C12N5/10, PC
	C12Q1/68,	
	PC	G01N33/53, G01N33/574//C12P21/02, C12P21/08, C12N15/00, A61K37/02,
	PC	A61K37/04,
	PC	C12N5/00
	CC	Compound and method for remedy and diagnosis of lung cancer PH
key		Location/Qualifiers

```

FEATURES
  source      location/Qualifiers
1. -2294     /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"

```

Query Match	100.0%	Score 1296;	DB 6;	Length 2294;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1296;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 ATGGAGGCGCCGTCGCGCGCGGCGCTGCTCTGCGGCGTCTGCGGAGGAGATCCAAAGC 60
Db 67 ATGAGAGCGCCGTCGCGCGCGGCGCTCTCTCGCTCGGCGGAGGAGATCCAAAGC 126
QY 61 AGCAATGAACTTCATCAAGTCCATCGAAGCTGTAAGTCTTAAATGGAGGAACAATGTGTG 120
Db 127 AGCAATGAACTTCATCAAGTCCATCGAAGCTGTAAGTCTTAAATGGAGGAACAATGTGTG 186

OY	121	TCGAAACAAGTACTTCTCCAAACATCTCACTGGTGGCAACTGGCCCAAGAAATTGGAGGGCAG	180
Db	187	TCGACCAAGTACTTCTCCAACTTCACTGTGCACTGGCCCAAGAAATTGGAGGGCAG	246
OY	181	CACGTGAAATAGATTAAGTCAAAACCTGCTATAGGGGAAATGGTCACTTTTACCGAGGA	240
Db	247	CACGTGAAATAGATTAAGTCAAAACCTGCTATAGGGGAAATGGTCACTTTTACCGAGGA	306
OY	241	AAGGCCAGCACTGACACCATGGGCGGCGCTTGCTGCTGGAACTGTGCACGTGTCTT	300
Db	307	AAGGCCAGCACTGACACCATGGGCGGCGCTTGCTGCTGGAACTGTGCACGTGTCTT	366
OY	301	CAGCAAAAGTACCTATGCCCAAGATCTGATGTCTTTCAAGCTGGGCTGGGGAAACATAT	360
Db	367	CAGCAAAAGTACCTATGCCCAAGATCTGATGTCTTTCAAGCTGGGCTGGGGAAACATAT	426
OY	361	TACGTGCAAGAACCCAGACAACCCGAGGAGACCCCTGGTGTATGTGACAGTGGGCTTAAG	420
Db	427	TACGTGCAAGAACCCAGACAACCCGAGGAGACCCCTGGTGTATGTGACAGTGGGCTTAAG	486
OY	421	CCGCTTGTCCAAGAGTGCATGTGTGCATGCTGCGAGATGGAAAAAGCCCTCTCTCT	480
Db	487	CCGCTTGTCCAAGAGTGCATGTGTGCATGCTGCGAGATGGAAAAAGCCCTCTCTCT	546
OY	481	CCAGAAAGATTTAAATTTCAGTGTGGCCAAAAGCTGTAGGCCCCGCTTTAAATTT	540
Db	547	CCAGAAAGATTTAAATTTCAGTGTGGCCAAAAGCTGTAGGCCCCGCTTTAAATTT	606
OY	541	GGGGGAGAAATCACACCATCTGAGAAACAGCCCTGTGGGCACTTACAGAGAGCAC	600
Db	607	GGGGGAGAAATCACACCATCTGAGAAACAGCCCTGTGGGCACTTACAGAGAGCAC	666
OY	601	CGGGGGGGGCTCTGTCACTACGTGTGTGTGAGGAGCAGCTCATCAGCCCTGTGGGTATC	660
Db	667	CGGGGGGGGCTCTGTCACTACGTGTGTGTGAGGAGCAGCTCATCAGCCCTGTGGGTATC	726
OY	661	AGCCGCACACACTGTCTCATTTGATTTACCCAAAGAGAGACTCATCTGTCTACCTGGGT	720
Db	727	AGCCGCACACACTGTCTCATTTGATTTACCCAAAGAGAGACTCATCTGTCTACCTGGGT	786
OY	721	CGCTCAAGGCTTAACTCCAAACAGCCAAAGGGGAGTGAAGTTTGAGGTGSAAAACCTATC	780
Db	787	CGCTCAAGGCTTAACTCCAAACAGCCAAAGGGGAGTGAAGTTTGAGGTGSAAAACCTATC	846
OY	781	CTACACAAGAGCTACAGCGCTGACACGCTTGCTTACCAACAACGATTTGCTTGCTGAAG	840
Db	847	CTACACAAGAGCTACAGCGCTGACACGCTTGCTTACCAACAACGATTTGCTTGCTGAAG	906
OY	841	ATCCGTTTCCAAGAGGAGGAGGTGTGGCAGCCATCCCGAGACTATACAGCCATCTGCCTG	900
Db	907	ATCCGTTTCCAAGAGGAGGAGGTGTGGCAGCCATCCCGAGACTATACAGCCATCTGCCTG	966
OY	901	CCCTCGAATGTATACGATTCGCCAGTTTGGACAAGCTGTAGATCATCGCTTTGGAAAA	960
Db	967	CCCTCGAATGTATACGATTCGCCAGTTTGGACAAGCTGTAGATCATCGCTTTGGAAAA	1026
OY	961	GAGAAATTTCAACGACTATCTTATCCGGAGAGCTGAAAAATGAATGTTGTGTAACCTATT	1020
Db	1027	GAGAAATTTCAACGACTATCTTATCCGGAGAGCTGAAAAATGAATGTTGTGTAACCTATT	1086
OY	1021	TCCCAACCGGAGGTGTACAGACGCCCACTACTACGGCTTGTGAAGTCAACACCAAAATCTG	1080
Db	1087	TCCCAACCGGAGGTGTGTACAGACGCCCACTACTACGGCTTGTGAAGTCAACACCAAAATCTG	1146
OY	1081	TGTGCTGTCTGACCCACAGTGGAAAAACGATTTCTGTCCAGGAGAGCTAGAGGGAGACCCCTC	1206
Db	1147	TGTGCTGTCTGACCCACAGTGGAAAAACGATTTCTGTCCAGGAGAGCTAGAGGGAGACCCCTC	1206
OY	1141	GTCGTGTTCCCTCCAAAGGCGCATGACTTTTGAAGTTTGAAGCTGGGCGCTGTGAATGT	1200
Db	1207	GTCGTGTTCCCTCCAAAGGCGCATGACTTTTGAAGTTTGAAGCTGGGCGCTGTGAATGT	1266
OY	1201	GCCTGTAAAGACAAAGCAGGCGTCTACAGAGACTTACACATTCTTAACCTGGATCCG	1260

Db 1267 GCCCGAAGGACACCGCGCTCTACAGAGTCTACACTTTTACCTTGATCCGC 1326
Qy 1261 AGTCACACCAAGAGAGAGATGCGCTGGCGCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGATGCGCTGGCGCTCTGA 1362

RESULT 11
HOMUKA
LOCUS 2294 bp mRNA linear PRI 17-FEB-1998
DEFINITION Homo sapiens mRNA for pro-urokinase precursor, complete cds.
ACCESSION D00244
VERSION D00244.1 GI:220138
KEYWORDS pro-urokinase precursor; pro-urokinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Verde, P., Stoppelli, M.P., Galeffi, P., Di Nocera, P. and Blasi, P.
TITLE Identification and primary sequence of an unspliced human urokinase
poly(A)+ RNA
Proc. Natl. Acad. Sci. U.S.A. 81 (15), 4727-4731 (1984)
JOURNAL 84:727-706
MEDLINE 6589620
PUBMED 6589620
REFERENCE
AUTHORS Nagai, M., Hiramatsu, R., Kaneda, T., Hayasuke, N., Arimura, H.,
Nishida, M. and Suyama, T.
TITLE Molecular cloning of cDNA coding for human preprourokinase
JOURNAL 86:56954
MEDLINE 2415429
PUBMED 2415429
REFERENCE
AUTHORS Hibino, Y., Miyake, T., Kobayashi, Y., Ohmori, M., Miki, T.,
Matsunoto, R., Numao, N. and Kondo, K.
TITLE Enhanced expression of human pro-urokinase cDNA in Escherichia coli
JOURNAL 52, 329-336 (1988)
FEATURES
SOURCE Location/Qualifiers
1..2294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pEB3, pKTU6, pKYU21"
/issue_type="kidney"
/note="2 bp upstream of PstI site"
67..1362
/codon_start=1
/product="pro-urokinase precursor"
/protein_id="BAA00175.1"
/db_xref="GI:220139"
/translation="MRALLARLLCVLVVSDSKSNELHQPENCDLNGTGVSNKY
FSNIHMCNCPKFGQHCIEDKSTCYEGHWRGKASDITMGRCPLPMNATYIQC
TYAHRSQDALQIGKHNYCRNDNRBRPVCYVGLKLNOCWHDADCKRSP
PEELKQCGGCKTTPRPKITGCEFTTENQWPAATYRHRGSAVTYVGGSLISFC
VIAATCFIDYPRKEDIYVLRSLNASTQEMKFEVENLLIHKDYSADTLAHNDI
ALKINISRCRCAQPSRTIQTICLPSMYNDPOFGTCEITGKEKNSTYLYLPEOLKM
TVKLISHRECOQPHYYSSEVITKMLCAADPOWKTDSCGDSGGPLVCSLQGRMTLTG
IVSWGSCALKDKRPYTVRSHPLPWRSHTEKNGIAL"

sig_peptide
mat_peptide
127..1359
/product="pro-urokinase"
708
/citation=[1]
/replace="g"
843
/citation=[2]
/replace="c"
1065
/citation=[2]
/replace="g"
1146
/citation=[1]

conflict
conflict
conflict

conflict /replace="a"
1161 /citation=[1]
/replace="c"
1164 /citation=[1]
/replace="a"
2259..2264
polym_signal
polym_signal
2272..2277
ORIGIN
Query Match 100.0%; Score 1296; DB 9; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAGCCCTGCTGCGCGCTGCTCTCTGCTCTGCTGAGCACTCCAAAGC 60
Db 67 ATGAGAGCCCTGCTGCGCGCTGCTCTCTGCTCTGCTGAGCACTCCAAAGC 126
Qy 61 AGCAATGAATCTCATCAATGATTCATCGAACTGTGACTGTCTAAATGAGAAACATGTGTG 120
Db 127 AGCAATGAATCTCATCAATGATTCATCGAACTGTGACTGTCTAAATGAGAAACATGTGTG 186
Qy 121 TCCAAAGATGATCTTCCCAACATTCATGCTGAGCACTGCCCAAGAAATTCGAGAGGAG 180
Db 187 TCCAAAGATGATCTTCCCAACATTCATGCTGAGCACTGCCCAAGAAATTCGAGAGGAG 246
Qy 181 CACTGTGAATATGATTAAGTCAAAAACCTGCTATGAGGGAAATGTCACCTTTTACCGAGGA 240
Db 247 CACTGTGAATATGATTAAGTCAAAAACCTGCTATGAGGGAAATGTCACCTTTTACCGAGGA 306
Qy 241 AAGGCGAGACATGACCAATGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 307 AAGGCGAGACATGACCAATGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
Qy 301 CAGCAAGATGACATGACCCCAAGATCTGATGCTCTTCACTGAGGCTGAGGAAACATTAAT 360
Db 367 CAGCAAGATGACATGACCCCAAGATCTGATGCTCTTCACTGAGGCTGAGGAAACATTAAT 426
Qy 361 TACTGCAAGAACCCGAGAACCGAGGCGACCTGCTGCTATGTCAGAGTGGGCTTAAG 420
Db 427 TACTGCAAGAACCCGAGAACCGAGGCGACCTGCTGCTATGTCAGAGTGGGCTTAAG 486
Qy 421 CGGCTTGTCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 487 CGGCTTGTCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
Qy 481 CCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
Db 547 CCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 606
Qy 541 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Db 607 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 666
Qy 601 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
Db 667 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 726
Qy 661 AGGCGCACACTGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db 727 AGGCGCACACTGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 786
Qy 721 GGGTCAAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
Db 787 GGGTCAAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 846
Qy 781 CTACACAAGGACTACAGGCTGACACGCTTGTCTACACCAAGACATGCTGCTGCTGCTGCTGCTGCT 840
Db 847 CTACACAAGGACTACAGGCTGACACGCTTGTCTACACCAAGACATGCTGCTGCTGCTGCTGCTGCT 906
Qy 841 ATCCGTTCCAAAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

Db 787 CGCTCAAGGCTTAACCTCCACACGCAAGGGGAGATGAAATTGGAGCTGAAAACTTAATC 846
Qy 781 CTACACAGGACTTACAGCGGCTGACACGCTTGTCTACCAACAGCAATTCCTTGTGAG 840
Db 847 CTACACAGGACTTACAGCGGCTGACACGCTTGTCTACCAACAGCAATTCCTTGTGAG 906
Qy 841 ATCCGTCACAGAGGAGGAGGTGTGCGAGCCATCCCGAGCTATACACACATCTGCTG 900
Db 907 ATCCGTCACAGAGGAGGAGGTGTGCGAGCCATCCCGAGCTATACACACATCTGCTG 966
Qy 901 CCCTGATGATATACGATGCCAGTTTGGACAAAGCTGTGAGATCACTGCTTTGAAAA 960
Db 967 CCCTGATGATATACGATGCCAGTTTGGACAAAGCTGTGAGATCACTGCTTTGAAAA 1026
Qy 961 GAGATTTTCAACGCTATCTCTATCCGAGAGAGCTGAAAAATGACTGTGTGAACTGATT 1020
Db 1027 GAGATTTTCAACGCTATCTCTATCCGAGAGAGCTGAAAAATGACTGTGTGAACTGATT 1086
Qy 1021 TCCCAACGAGAGTGTACAGAGCCCACTACAGGCTGTGAAGTCAACCAAAATGCTG 1080
Db 1087 TCCCAACGAGAGTGTACAGAGCCCACTACAGGCTGTGAAGTCAACCAAAATGCTG 1146
Qy 1081 TGTGTGTGACCCCAAGTGTGAAAAAGATTCTGCGAGGAGACTCAAGGGAGCCCTC 1140
Db 1147 TGTGTGTGACCCCAAGTGTGAAAAAGATTCTGCGAGGAGACTCAAGGGAGCCCTC 1206
Qy 1141 GTCTGTTCCTTCCAAAGGCGCATGACTTGTGACTGGAATTGTGAGCTGGGGCCGTGATG 1200
Db 1207 GTCTGTTCCTTCCAAAGGCGCATGACTTGTGACTGGAATTGTGAGCTGGGGCCGTGATG 1266
Qy 1201 GCCCTGAAGGACAAAGCCAGGCGTGTACACAGAGTCTTCACTTTACCTGGATCCG 1260
Db 1267 GCCCTGAAGGACAAAGCCAGGCGTGTACACAGAGTCTTCACTTTACCTGGATCCG 1326
Qy 1261 AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGA 1236
Db 1327 AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGA 1362

RESULT 13
A35395
LOCUS A35395 2296 bp RNA linear PART 11-DEC-1996
DEFINITION H.sapiens u-PA cDNA sequence.
ACCESSION A35395
VERSION A35395.1 GI:1926844
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2296)
AUTHORS Meyhack,B., Heim,U. and Buerigl,R.
TITLE Process for the production of proteins
JOURNAL Patent: EP 0288435-A 2 26-OCT-1988;
CIBA-GEIGY AG
FEATURES
source location/Qualifiers
1..2296
/organism="Homo sapiens"
/mol_type="unassigned RNA"
/db_xref="taxon:9606"
70..1365
/gene="u-PA"
70..1365
/gene="u-PA"
/codon_start=1
/protein_id="CAA02215.1"
/db_xref="GI:1926845"
/translation="MRALLARLLCYLVIVSDSGSGNEHQPVSNCDCINGSTCVSNKY
FSYIHWKNCPPKRGQGHCEIDSKYCTIEBNGHFYRGKSTDTMGSPCLPMSATVLOQ
TYHAHSDALQJLGLKHNKYNRNDNRKRWCVQVGLRPLVECHWDCADGKXKSP
PELRFQCGQKTLRPFKXIGBFTTINQPFPAIYRHRGSGVTYVCGSLISPCW
VTSATHCFLDPKEDYIVYLGSRSLNSNTQEMKFEVENLILHSDYASDTLAHNDI

ORIGIN

Query Match 99.9%; Score 1294.4; DB 6; Length 2296;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ALLIKRSKEGRCAQPSRTIQTICLPSMYNDPQGTSCETITGPKENSTDYLYPEQLKM
TVYKLISHRECOQPHYSGSEVTTKMLCAADPQMKTDSCQGBSGPLVSLGRMTLTG
IVSMRGKALXKDKPGVYTRVSHFLPWISHTKEENGALH"
Qy 1 ATGAGAGCCCTGCTGGCGGCGCTTCTCTGCGCTCTGCTGTGAGGAGCACTCCAAAGC 60
Db 70 ATGAGAGCCCTGCTGGCGGCGCTTCTCTGCGCTCTGCTGTGAGGAGCACTCCAAAGC 129
Qy 61 AGCAATGAACCTTCATCAAGTTCATCGAACTGTGACTGTAAATGAGAGCAATGTGTG 120
Db 130 AGCAATGAACCTTCATCAAGTTCATCGAACTGTGACTGTAAATGAGAGCAATGTGTG 189
Qy 121 TCCAAAGTACTTCTCCCAAGTTCATCGAGTGTGCACTGCCCCAAGAAATTTGGAGGCGAG 180
Db 190 TCCAAAGTACTTCTCCCAAGTTCATCGAGTGTGCACTGCCCCAAGAAATTTGGAGGCGAG 249
Qy 181 CACTGTGAATAGATTAAGTCAAAAACCTGCTATGAGGGAAATGTCACCTTTACCGAGA 240
Db 250 CACTGTGAATAGATTAAGTCAAAAACCTGCTATGAGGGAAATGTCACCTTTACCGAGA 309
Qy 241 AAGGCGAGACTGACACCATGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 310 AAGGCGAGACTGACACCATGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
Qy 301 CAGCAAAAGTACCAAGTCCAGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 370 CAGCAAAAGTACCAAGTCCAGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
Qy 361 TACTCAGAGAACCCAGACACCGAGGCGACCTGTGCTATGTGACAGTGGGCTTAAG 420
Db 430 TACTCAGAGAACCCAGACACCGAGGCGACCTGTGCTATGTGACAGTGGGCTTAAG 489
Qy 421 CGGCTTGTCCAAAGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 490 CGGCTTGTCCAAAGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
Qy 481 CCAGAAAGATTAATAATTTCAGTGTGCGCAAAAGACTGTGAGGCGCCGCTTTAAGATTA 540
Db 550 CCAGAAAGATTAATAATTTCAGTGTGCGCAAAAGACTGTGAGGCGCCGCTTTAAGATTA 609
Qy 541 GGGGAGAAATTCACCACTATGAGAACCAAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 610 GGGGAGAAATTCACCACTATGAGAACCAAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 669
Qy 601 CGGCGGGGCTGTGTCACTAGTGTGAGGAGGAGCTCAATCAAGCCTTGTGCTGCTGATG 660
Db 670 CGGCGGGGCTGTGTCACTAGTGTGAGGAGGAGCTCAATCAAGCCTTGTGCTGCTGATG 729
Qy 661 AGCGGCACACACTGCTTCATGATTAACCAAGAGAGAGACTACATGCTTACCTGGGT 720
Db 730 AGCGGCACACACTGCTTCATGATTAACCAAGAGAGAGACTACATGCTTACCTGGGT 789
Qy 721 CGCTCAAGGCTTAATCTCAACAGCGAGGAGATGAAGTTTGAAGTGGAAAACTTAATC 780
Db 790 CGCTCAAGGCTTAATCTCAACAGCGAGGAGATGAAGTTTGAAGTGGAAAACTTAATC 849
Qy 781 CTACACAGGACTTACAGCGGCTGACAGCTTGTCTACCAACAGCAATTCCTTGTGAG 840
Db 850 CTACACAGGACTTACAGCGGCTGACAGCTTGTCTACCAACAGCAATTCCTTGTGAG 909
Qy 841 ATCCGTTCCAAAGAGGCGAGGTGTGCGAGCCATCCCGAGCTATACAGACATCTGCTG 900
Db 910 ATCCGTTCCAAAGAGGCGAGGTGTGCGAGCCATCCCGAGCTATACAGACATCTGCTG 969
Qy 901 CCCTGATGATATACGATCCCGAGTTTGGCAAGCTGTGAGATCACTGCTTTGAAAA 960
Db 970 CCCTGATGATATACGATCCCGAGTTTGGCAAGCTGTGAGATCACTGCTTTGAAAA 1029

Oy		961	GAGAAATTCACCGACTATTCTCATCTCCGAGAGCGTGAATAAGA	CTGTGTGAAGTGATT	1022
Db		1030	GAGAAATTCACCGACTATTCTCATCTCCGAGAGCGTGAATAAGA	CTGTGTGAAGTGATT	10899
Oy		1021	TCCCAACCGGGAGTGTACAGCAGCCCCCATCTACTACAGGCCT	CTGAAGTCAACCACCAAATGCTG	10808
Db		1090	TCCCAACCGGGAGTGTACAGCAGCCCCCATCTACTACAGGCCT	CTGAAGTCAACCACCAAATGCTG	11449
Oy		1081	TGTGTGTCTGACCCCAAGTGGAAAAACAATTCTCTGCCAGGGAG	ACTCAGGGGAGACCCCTC	11400
Db		1150	TGTGTGTCTGACCCCAAGTGGAAAAACAATTCTCTGCCAGGGAG	ACTCAGGGGAGACCCCTC	12099
Oy		1141	GTCGTCTCCCCTCCAAAGCGCCGATGACTTTGA CTGGAAATTG	AAGTGTGGAGTCGGGGCGGTGGAAATG	12000
Db		1210	GTCGTCTCCCCTCCAAAGCGCCGATGACTTTGA CTGGAAATTG	AAGTGTGGAGTCGGGGCGGTGGAAATG	12699
Oy		1201	GCCCTGAAGAACAAGCCAGCGGCTCTACACGAGAAGTCTCAC	ATTCTTACTTCCTCTGATCCGC	12600
Db		1270	GCCCTGAAGAACAAGCCAGCGGCTCTACACGAGAAGTCTCAC	ATTCTTACTTCCTCTGATCCGC	13299
Oy		1261	AGTCACACCAAGAAAGAGAATGGCTGGCCCTCTTGA	1296	
Db		1330	AGTCACACCAAGAAAGAGAATGGCTGGCCCTCTTGA	1365	
RESULT_14					
BC013575					
LOCUS					
DEFINITION					
Homo sapiens plasminogen activator, urokinase, mRNA (CDNA clone MGC:9223 IMAGE:3890980), complete cds.					
ACCESSION					
BC013575.2 GI:34194099					
VERSION					
MGC.					
KEYWORDS					
MGC.					
SOURCE					
Homo sapiens (human)					
ORGANISM					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE					
AUTHORS					
Strauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shanem,C.M., Schuler,G.D., Alschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan H., Moore,T., Max,S.I., Wang,J., Heieh,F., DiChienko,L., Marusina,K., Farmer,A.A., Riddin,G.W., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schneez,T.E., Brownstein,M.J., Udell,T.B., Toohiyuki,S., Canciani,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Boask,S.A., Meban,P.J., McKernan,K.U., Malek,U.A., Gunaratne,P.H., Richards,S., Wooley,K.C., Hale,S., Garcia,A.M., Gay,L.U.J., Hulys,S.W., Villalton,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Keltzman,M., Madan,A.C., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Buterfield,Y.S., Krzywinski,W.I., Skalska,J., Smalls,D.E., Scherch,I., Schein,J.B., Jones,S.J. and Marx,M.A.					
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences					
JOURNAL					
MEDLINE					
PUBMED					
22388257 Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)					
REFERENCE					
TITLE					
2 (bases 1 to 2333)					
JOURNAL					
AUTHORS					
Submitted (09-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomes Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
REMARK					
NIH-MGC Project URL: http://mgc.nci.nih.gov					
On Aug 25, 2003 this sequence version replaced gi:15488888.					
COMMENT					
Contact: MGC help desk					
Email: cgapbs-re@mail.nih.gov					
Tissue Procurement: DCTD/DTP/Gazdar					

DNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAK P1.6.E. Consortium LINL at: <http://image.lnl.gov>
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505862.

```
FEATURES
    source
        1..2333
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="MGC:9223 IMAGE:3890980"
            /tissue_type="lung, carcinoma, large cell undifferentiated."
            /clone_id="NIH_MGC_69"
            /lab_host="DH10B"
            /note="vector: pCMV-SPORT6"
        1..2333
            /gene="PLAU"
            /note="synonyms: URK, UPA"
            /db_xref="LocustID:5328"
            /db_xref="MIM:191840"
            93..1388
                /codon_start=1
                /product="plasminogen activator, urokinase"
                /protein_id="AAH13575.1"
                /db_xref="GI:15488889"
                /db_xref="LocustID:5328"
                /translation="MRALLRLILCVLVSDSKSNELHVPNSDCLNGCTCSNRYFSNHRCSPKRTGGQHCETIDKSCTCYEGSHYRGCASTDWGRPCLPWNASTVLQQTYAHNRDALQLGIGKNAYCRNDNRNPVCYOVGLP/VDECMVHDCADGKKPSRPEELKFCQGCKTLRPRIKIIGESTITENQPFALIRRRGSVTYCGSLISPCMVISATCFDIYPKKEEDYIVYLGRSLNSNGEMKEVENULIHKDYASDTLAHNIDIALKIKRSRECCOQPHYTGSEVTTKMLCADPQKTDSCQGDSGGPLVCSLQGMILLTGTVAKLISHRECQDPHYTGSEVTTKMLCADPQKTDSCQGDSGGPLVCSLQGMILLTGTIVSMGRCALDKDRGYVTRVSHFLPWIRSHTKENGIAL"
                300..545
                    /note="Kringletype Region: Kringletype domain have been found in plasminogen, hepatocyte growth factor, prothrombin, and apolipoprotein A. Structure is disulfide-rich, nearly all-beta"
                    /db_xref="CDD:pfam0051"
                    624..1349
                        /note="TIYP SPC Region: Trypsin-like serine protease"
                        /db_xref="CDD:smart0020"
```

	Query Match	Best Local Similarity	Score	Pred. No.	Mismatches	Indels	Gaps
Origins	Matches	1295	Conservative	0	1	0	0
misc_feature	1	ATGAGAGCCCTGCTGGCGGCCCTGCTTCCTGCGTCTGTTGCTGTAAGCGACTCCAAAGC	60	99.9%	1294.4	DB 9	Length 2333
Db	93	ATGAGAGCCCTGCTGGCGGCCCTGCTTCCTGCGTCTGTTGCTGTAAGCGACTCCAAAGC	152				
Oy	61	AGCAATGAACCTCATCAAGTTTCATGCACTGTCAGCTGTAATGAGAGAACAATGTC	120				
Db	153	AGCAATGAACCTCATCAAGTTTCATGCACTGTCAGCTGTAATGAGAGAACAATGTC	212				
Oy	121	TCCAACAAGTACTTTCACCAATTCACCTGTCGACACTCCCAAGAAATTGCGAGGCGAG	180				
Db	213	TCCAACAAGTACTTTCACCAATTCACCTGTCGACACTCCCAAGAAATTGCGAGGCGAG	272				
Oy	181	CACGTGAATAGATAAGTAAAAAAGCTGCTATGAGGGGAATGTCACCTTTTACCGAGA	240				

Db	273	CACGTGAAATPAGATTAAGTAAATAAACCCTGCTATAGAGGGAAATGGTCACTTTACCGAGGA	333
QY	241	AAGGCCACGACCTGACACCAATGAGGCGCGGCCCTGCTGCCTTGAAACTGTGCACATGCTCTT	300
Db	333	AAGGCCACGACCTGACACCAATGAGGCGCGGCCCTGCTGCCTTGAAACTGTGCACATGCTCTT	392
QY	301	CAGCAACGTAACATGCGCCACAGATCTGATGCTTTTGAAGCTGGGCTGGGGAAACATAAT	366
Db	393	CAGCAACGTAACATGCGCCACAGATCTGATGCTTTTGAAGCTGGGCTGGGGAAACATAAT	452
QY	361	TACTGCAGGAACCCAGACAAACCGAGGCGACCTTGATGCTATGTGCAAGTGGGCTTAAG	420
Db	453	TACTGCAGGAACCCAGACAAACCGAGGCGACCTTGATGCTATGTGCAAGTGGGCTTAAG	512
QY	421	CCGCTGTCCAAAGTGTCATGATGCAATGACTGGGCGACATGGAATAAAAGCCCTCTCTCT	480
Db	513	CCGCTGTGTCCAAAGTGTCATGATGCAATGACTGGGCGACATGGAATAAAAGCCCTCTCTCT	572
QY	481	CCAGAAAGATTTAAATTTTCAGTGTGGCCAAAAGACTGTGAGGCGCCGCTTTAAGATTAAT	540
Db	573	CCAGAAAGATTTAAATTTTCAGTGTGGCCAAAAGACTGTGAGGCGCCGCTTTAAGATTAAT	632
QY	541	GGGGGGAATTTCAACCAATCGAGAACCAAGCCTTGTTTGGGCGACATTAACAGAGGAC	600
Db	633	GGGGGGAATTTCAACCAATCGAGAACCAAGCCTTGTTTGGGCGACATTAACAGAGGAC	692
QY	601	CGGGGGGGGCTCTGTCACTACGATGATGAGGAGAGCCTCATCAAGCCTTGCTGGGATG	660
Db	693	CGGGGGGGGCTCTGTCACTACGATGATGAGGAGAGCCTCATCAAGCCTTGCTGGGATG	752
QY	661	AAGGCCACACATGCTTCATTGATTAACCAAGAAGAGAGATCAATCGTCTACCTGGGT	720
Db	753	AAGGCCACACATGCTTCATTGATTAACCAAGAAGAGAGATCAATCGTCTACCTGGGT	812
QY	721	CGCTCAAGGCTTAATCTCCAAACAGCAAGGGGAGATGAAGTTTGAAGTGAATAACCTAATC	780
Db	813	CGCTCAAGGCTTAATCTCCAAACAGCAAGGGGAGATGAAGTTTGAAGTGAATAACCTAATC	872
QY	781	CTACACAGAAGCTTACAGCGCTGACACGCTTGCTCAACCAACAGCATTTGCTTGGAAG	840
Db	873	CTACACAGAAGCTTACAGCGCTGACACGCTTGCTCAACCAACAGCATTTGCTTGGAAG	932
QY	841	ATCCGTTTCAAGGAGGGGAGGTGTGGGAGCCATCCCGGACTTAACAGCATCTGCTG	900
Db	933	ATCCGTTTCAAGGAGGGGAGGTGTGGGAGCCATCCCGGACTTAACAGCATCTGCTG	992
QY	901	CCCTCGATGATTAACCATCCAGTTTGGCAACAAGCTGAGATCACTCGGCTTTGGAAA	960
Db	993	CCCTCGATGATTAACCATCCAGTTTGGCAACAAGCTGAGATCACTCGGCTTTGGAAA	1052
QY	961	GAGAAATTCATCCGACTATCTTATCCGAGCAGCTGAAAATGATCTGTTTGAAGCTGATT	1020
Db	1053	GAGAAATTCATCCGACTATCTTATCCGAGCAGCTGAAAATGATCTGTTTGAAGCTGATT	1112
QY	1021	TCCCAACCGGAGTGTACAGAGCCGCCACTCTACGAGCTCGAAGTCAACCAAAATGCTG	1080
Db	1113	TCCCAACCGGAGTGTACAGAGCCGCCACTCTACGAGCTCGAAGTCAACCAAAATGCTG	1172
QY	1081	TGTGCTGCTGACCCACAGTGAATAACAGATTCTGTGCAAGGAGACTCAAGGAGGACCCCTC	1140
Db	1173	TGTGCTGCTGACCCACAGTGAATAACAGATTCTGTGCAAGGAGACTCAAGGAGGACCCCTC	1232
QY	1141	GTCCTGTTCCCTCCAAAGGCGCGATGACTTTGATCTGGAATTGTGAGCTGGGGCCGTGGATGT	1200
Db	1233	GTCCTGTTCCCTCCAAAGGCGCGATGACTTTGATCTGGAATTGTGAGCTGGGGCCGTGGATGT	1292
QY	1201	GCCCTGAAGAGCAAGCCAGGCGCTTACAAGAGTCTCAACTTCTTAACCTGTAACTCCG	1260
Db	1293	GCCCTGAAGAGCAAGCCAGGCGCTTACAAGAGTCTCAACTTCTTAACCTGTAACTCCG	1352
QY	1261	AGTCACACCAAGGAGGAATGGCTGGGCCCTCTCTGA	1296
Db	1353	AGTCACACCAAGGAGGAATGGCTGGGCCCTCTCTGA	1388

[illegible]

QY 301 CAGCAAGGATCCATGCCCCAGATCTGATCTTTCAGCTGGGCCCTGGGGAACAATAT 360
DB 301 CAGCAAGGATCCATGCCCCAGATCTGATCTTTCAGCTGGGCCCTGGGGAACAATAT 360
QY 361 TACTGAGGAACCCAGACCAACCGAGGCGAACCTGTGTCTATGTGAGTGGGCTTAAG 420
DB 361 TACTGAGGAACCCAGACCAACCGAGGCGAACCTGTGTCTATGTGAGTGGGCTTAAG 420
QY 421 CCGCTTGTCCAGAGTGCATGTGTGATGACTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
DB 421 CCGCTTGTCCAGAGTGCATGTGTGATGACTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
QY 481 CCAAGAAATTAATTTTCAGTGTGGCCAAAGACTGAGGCCCGCTTAAGATTAT 540
DB 481 CCAAGAAATTAATTTTCAGTGTGGCCAAAGACTGAGGCCCGCTTAAGATTAT 540
QY 541 GGGGAGAAATTCACCAATCGAGAACAGCCCTGTGTTGGGCCATCTACAGAGGAC 600
DB 541 GGGGAGAAATTCACCAATCGAGAACAGCCCTGTGTTGGGCCATCTACAGAGGAC 600
QY 601 CCGGGGGGCTCTGTCACTTACGTGTGTGAGGAGCAAGCTCATCAGCCCTGTGGTGATC 660
DB 601 CCGGGGGGCTCTGTCACTTACGTGTGTGAGGAGCAAGCTCATCAGCCCTGTGGTGATC 660
QY 661 AGGCGCACACACGCTTCAATGATTAACCCAAAGAGAGATGAATTGAGGTGAAACTATC 720
DB 661 AGGCGCACACACGCTTCAATGATTAACCCAAAGAGAGATGAATTGAGGTGAAACTATC 720
QY 721 CCGTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTGAGGTGAAACTATC 780
DB 721 CCGTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTGAGGTGAAACTATC 780
QY 781 CTACACAAGACTTACAGCGCTGACAGCTTGTCTACCAACAACATTTGCTTGAAG 840
DB 781 CTACACAAGACTTACAGCGCTGACAGCTTGTCTACCAACAACATTTGCTTGAAG 840
QY 841 ATCCGTTCCAGAGAGGGGAGAGTGTGGCGACCATCCCGGACTATACAGACCATCTGCTG 900
DB 841 ATCCGTTCCAGAGAGGGGAGAGTGTGGCGACCATCCCGGACTATACAGACCATCTGCTG 900
QY 901 CCTCGATGATTAACGATCCCGAGTTTGGCAACAAGCTGTGAGATCACTGCTTGGAAA 960
DB 901 CCTCGATGATTAACGATCCCGAGTTTGGCAACAAGCTGTGAGATCACTGCTTGGAAA 960
QY 961 GAGAAATTTACCGACTATCTTATCCGAGACAGCTGAAAATGACTGTTGAACTGATT 1020
DB 961 GAGAAATTTACCGACTATCTTATCCGAGACAGCTGAAAATGACTGTTGAACTGATT 1020
QY 1021 TCCCAACGGGAGAGTGTGAGCAGCCCACTATACGGCTCTGAAGTCAACACCAAAATGCTG 1080
DB 1021 TCCCAACGGGAGAGTGTGAGCAGCCCACTATACGGCTCTGAAGTCAACACCAAAATGCTG 1080
QY 1081 TGTGTGTGTGATCCCAAGTGTGAAACAGATTCTGTCCAGGGAGACTCAGGGGAGCCCTC 1140
DB 1081 TGTGTGTGTGATCCCAAGTGTGAAACAGATTCTGTCCAGGGAGACTCAGGGGAGCCCTC 1140
QY 1141 GTCTGTTCCTCCAGGCGGATGACTTTGATGGAATTGTGAGCTGGGGCCGTGATGT 1200
DB 1141 GTCTGTTCCTCCAGGCGGATGACTTTGATGGAATTGTGAGCTGGGGCCGTGATGT 1200
QY 1201 GCGCTGAAGAGCAAGCGGCTGTACAGAGTGTCAACATTTTAACTCTGATCCG 1260
DB 1201 GCGCTGAAGAGCAAGCGGCTGTACAGAGTGTCAACATTTTAACTCTGATCCG 1260
QY 1261 AGTCAACCAAGAGAGATGAGCTGAGCCCTCTGA 1296
DB 1261 AGTCAACCAAGAGAGATGAGCTGAGCCCTCTGA 1296

THIS PAGE BLANK (USPTO)

XX New isolated lung tumor polynucleotides, used to develop products for the
 PT treatment, prevention and monitoring the progression of lung cancer.

XX Claim 12; Page 115-116; 148pp; English.

XX The invention provides isolated human lung tumor nucleic acids and
 CC polypeptides. The polypeptides can be used for the treatment of lung
 CC cancer. The polypeptides and polynucleotides can be used to stimulate T
 CC cells or antigen presenting cells for use in the treatment of lung
 CC cancer. The polypeptides and monoclonal antibodies specific for the
 CC polypeptides can also be used to inhibit the development of lung cancer.
 CC Agents which bind the polypeptides can be used for detecting lung cancer
 CC and for monitoring the progression of lung cancer

XX Sequence 2281 BP; 549 A; 595 C; 604 G; 533 T; 0 U; 0 Other;

Query Match 100.0%; Score 1296; DB 2; Length 2281;
 Best Local Similarity 100.0%; Pred. No. 2.9e-267;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGCGCTGCTTCTGCGTCTGCTGCGCATCCCAAGGC 60
 DB ATGAGAGCCCTGCTGCGCGCTGCTTCTGCGTCTGCTGCGCATCCCAAGGC 126
 QY 61 AGCAATGAATCTATCAAGTTCATGCACTGCACTGCTTAATGAGAGACATGTG 120
 DB AGCAATGAATCTATCAAGTTCATGCACTGCACTGCTTAATGAGAGACATGTG 186
 QY 121 TCACAAGATCTTCTCAACATTCATGCTGCTGCACTGCCAAGAAATTCGAGGGCAG 180
 DB TCACAAGATCTTCTCAACATTCATGCTGCTGCACTGCCAAGAAATTCGAGGGCAG 246
 QY 181 CACTGTGAATATAGTAAAGTAAACCTGCTATGAGGGGAATGTCATTTTACGAGGA 240
 DB CACTGTGAATATAGTAAAGTAAACCTGCTATGAGGGGAATGTCATTTTACGAGGA 306
 QY 241 AAGGCGAGACTGACACCAATGCGCGCGCTGCTGCTGCACTGCTGCTT 300
 DB AAGGCGAGACTGACACCAATGCGCGCGCTGCTGCTGCACTGCTGCTT 366
 QY 301 CAGCAAGATGATCCATCCCAAGATGATGATGCTTCACTGCGCGCTGCGGAAACATTAAT 360
 DB CAGCAAGATGATCCATCCCAAGATGATGATGCTTCACTGCGCGCTGCGGAAACATTAAT 426
 QY 361 TACTGAGAGACCCAGACCAACCGAGGCGACCTGCTGCTATGAGAGTGGCTTAAG 420
 DB TACTGAGAGACCCAGACCAACCGAGGCGACCTGCTGCTATGAGAGTGGCTTAAG 486
 QY 421 CCGCTTGTCCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB CCGCTTGTCCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
 QY 481 CCAGAAAGATTAATTTTCAAGTGTGCGCAAAAGACTCTGAGGCGCGCTTAAAGTATT 540
 DB CCAGAAAGATTAATTTTCAAGTGTGCGCAAAAGACTCTGAGGCGCGCTTAAAGTATT 606
 QY 541 GGGGGAGAAATTCACCAATCGAGAACAGCCCTGCTGCTGCGGCTTAAAGAGGAC 600
 DB GGGGGAGAAATTCACCAATCGAGAACAGCCCTGCTGCTGCGGCTTAAAGAGGAC 666
 QY 601 CCGGGGGGCTTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 DB CCGGGGGGCTTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
 QY 661 AGGCGACACACTGCTTCAATGATTAACCAAGAGAGAGACTGCTGCTGCTGCTGCTGCT 720
 DB AGGCGACACACTGCTTCAATGATTAACCAAGAGAGAGACTGCTGCTGCTGCTGCTGCT 786
 QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAAGTTGAGGTGAGAAACCTTAATC 780
 DB CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAAGTTGAGGTGAGAAACCTTAATC 846

QY 781 CTACACAAGAGACTACAGCGCTGACAGCTTGTCTACACACAGACATTCGCTTGTGAAG 840
 DB CTACACAAGAGACTACAGCGCTGACAGCTTGTCTACACACAGACATTCGCTTGTGAAG 906
 QY 841 ATCCGTTCCAGAGAGGAGAGTGTGCGACCCATCCCGGACTATACAGACATTCGCTG 900
 DB ATCCGTTCCAGAGAGGAGAGTGTGCGACCCATCCCGGACTATACAGACATTCGCTG 966
 QY 901 CCTCGATATATACATATCCCAAGTTGGGACCAAGCTGTGAGATCACTGGCTTTGAGAAA 960
 DB CCTCGATATATACATATCCCAAGTTGGGACCAAGCTGTGAGATCACTGGCTTTGAGAAA 1026
 QY 961 GAGAAATTCACGACTATCTATTCGAGAGACGCTGAAATGACTGTTGAGAGTGAAT 1020
 DB GAGAAATTCACGACTATCTATTCGAGAGACGCTGAAATGACTGTTGAGAGTGAAT 1086
 QY 1021 TCCACCGGAGAGTGTGACAGAGCCCACTACAGGCTTGAAGTCAACCAAAATGCTG 1080
 DB TCCACCGGAGAGTGTGACAGAGCCCACTACAGGCTTGAAGTCAACCAAAATGCTG 1146
 QY 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTCCTGCGAGGAGACTCAGGGGAGCCCTTC 1140
 DB TGTGCTGTGACCCACAGTGAAGAAACAGATTCCTGCGAGGAGACTCAGGGGAGCCCTTC 1206
 QY 1141 GTCTGTTCCTCCAGAGCGCGAATGACTTGAATGAAATGAGTGGGGCGGTGATGT 1200
 DB GTCTGTTCCTCCAGAGCGCGAATGACTTGAATGAAATGAGTGGGGCGGTGATGT 1266
 QY 1201 GCCCTGAAGACAGAGCAGGCGTCTACAGAGTCTCACTTCTTAACCTGATCCGC 1260
 DB GCCCTGAAGACAGAGCAGGCGTCTACAGAGTCTCACTTCTTAACCTGATCCGC 1326
 QY 1261 AGTCACACAGAGAGAGATGAGTGGCTGCGCTCTGA 1296
 DB AGTCACACAGAGAGAGATGAGTGGCTGCGCTCTGA 1362

RESULT 3
 ID AAC65859 standard; cDNA; 2294 BP.
 XX
 AC AAC65859;
 XX
 DT 21-FEB-2001 (first entry)
 XX
 DE Human lung cancer-associated cDNA for contig 12.
 XX
 DE Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
 KW vaccine; detection; sg.
 XX
 OS Homo sapiens.
 OS
 PN WO20061612-A2.
 PD 19-OCT-2000.
 XX
 PF 03-APR-2000; 2000WO-US008896.
 XX
 PR 02-APR-1999; 99US-00285479.
 PR 17-DEC-1999; 99US-00466396.
 PR 30-DEC-1999; 99US-00476496.
 PR 10-JAN-2000; 2000US-00480884.
 PR 22-FEB-2000; 2000US-00510376.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Fan L;
 XX
 DR WPI; 2000-628399/60.
 XX
 PT Isolated polypeptide comprising an immunogenic portion of a lung tumor
 PT protein is used for detecting and monitoring progression of lung cancer
 PT in a patient.

XX Polynucleotides encoding lung tumor polypeptides, useful for treating
 PT lung cancer or stimulating an immune response.
 XX
 PS Example 3; Page 232-233; 374pp; English.
 CC The present invention describes human lung tumour proteins, Human lung
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant
 CC activities, and can be used in vaccine production. Compositions
 CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
 CC proteins, T cell populations, or antigen presenting cells that express
 CC the lung tumour proteins are useful for treating lung cancer or
 CC stimulating an immune response. ABL48959 to ABL49300 and ABL74946 to
 CC ABL75070 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 2294 BP; 549 A; 596 C; 604 G; 545 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1296; DB 6; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 2.9e-267;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ATGAGAGCCCTGTGCGCGCTTCTCTGCTCTGCTGCTGAGCACTCCAAAGGC 60
 Db 67 ATGAGAGCCCTGTGCGCGCTTCTCTGCTCTGCTGCTGAGCACTCCAAAGGC 126
 Oy 61 AGCAATGAATCTTCATCAAGTTCCTCATGAACTGTGACTGTCTTAATGAGAAACATGTGTG 120
 Db 127 ACCAATGAATCTTCATCAAGTTCCTCATGAACTGTGACTGTCTTAATGAGAAACATGTGTG 186
 Oy 121 TCCCAACAGTACTCTCCCAACATTCACCTGTCGAACTGCGCCAAAGAAATTCGAGGGCAG 180
 Db 187 TCCCAACAGTACTCTCCCAACATTCACCTGTCGAACTGCGCCAAAGAAATTCGAGGGCAG 246
 Oy 181 CACTGTGAATAGATAGTCAAAAACCTGCTATGAGGGAGATGCTCACTTTTACCAGAGA 240
 Db 247 CACTGTGAATAGATAGTCAAAAACCTGCTATGAGGGAGATGCTCACTTTTACCAGAGA 306
 Oy 241 AAGGCGAGCTGACACCACTGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 Db 307 AAGGCGAGCTGACACCACTGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
 Oy 301 CAGCAAAAGCTACCACTGCGCCCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 Db 367 CAGCAAAAGCTACCACTGCGCCCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
 Oy 361 TACTGAGAGAACCCAGACCAACCGAGAGCACTTGTGCTATGTGAGGTGGGCTTAAG 420
 Db 427 TACTGAGAGAACCCAGACCAACCGAGAGCACTTGTGCTATGTGAGGTGGGCTTAAG 486
 Oy 421 CCGCTTGTCCAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 487 CCGCTTGTCCAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
 Oy 481 CCAAGAAATTTAAATTTTAAAGTGTGCGCAAAAGATCTCTAGAGGCGCGCTTTAAGATTATT 540
 Db 547 CCAAGAAATTTAAATTTTAAAGTGTGCGCAAAAGATCTCTAGAGGCGCGCTTTAAGATTATT 606
 Oy 541 GGGGGGAATTTACCAACCATCGAGAACAGCCCTGTGTTGGCGGCATCTTACAGAGAGCAC 600
 Db 607 GGGGGGAATTTACCAACCATCGAGAACAGCCCTGTGTTGGCGGCATCTTACAGAGAGCAC 666
 Oy 601 GGGGGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 667 GGGGGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
 Oy 661 AAGGCGACACACTGCTTCAATTGATTATCCAAAGAAAGAGAGACATCTTACCTGAGGT 720
 Db 727 AAGGCGACACACTGCTTCAATTGATTATCCAAAGAAAGAGAGACATCTTACCTGAGGT 786
 Oy 721 CGCTCAAGGCTTAACCTCAACACGCAAGGGAGATGAAGTTGGAGGTGAAAACTTAATC 780
 Db 787 CGCTCAAGGCTTAACCTCAACACGCAAGGGAGATGAAGTTGGAGGTGAAAACTTAATC 846

Oy 781 CTACACAAGAGTACTACAGCGCTGACACGCTTGTCTACACACAGACATGCTGCTGAGG 840
 Db 847 CTACACAAGAGTACTACAGCGCTGACACGCTTGTCTACACACAGACATGCTGCTGAGG 906
 Oy 841 ATCCGCTCCAAAGAGGAGAGTGTGCGCAGCCATCCCGGACTATACAGACATCTGCTG 900
 Db 907 ATCCGCTCCAAAGAGGAGAGTGTGCGCAGCCATCCCGGACTATACAGACATCTGCTG 966
 Oy 901 CCGTGTATATPACGATCCCAAGTTTGGCAAGAGTGTGAGATCACTGGCTTTGAAAA 960
 Db 967 CCGTGTATATPACGATCCCAAGTTTGGCAAGAGTGTGAGATCACTGGCTTTGAAAA 1026
 Oy 961 GGAATTTACCAAGATATCTATATCCGAGCAGCTGAAAATGACGTTGTGAAGCTGATT 1020
 Db 1027 GGAATTTACCAAGATATCTATATCCGAGCAGCTGAAAATGACGTTGTGAAGCTGATT 1086
 Oy 1021 TCCCAACGGAGATGTACAGACGCCCACTACTACGCTGTAAGTCAACCAAAATGCTG 1080
 Db 1087 TCCCAACGGAGATGTACAGACGCCCACTACTACGCTGTAAGTCAACCAAAATGCTG 1146
 Oy 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTCCTGCGAGGAGACTCAGGGGGACCCCTC 1140
 Db 1147 TGTGCTGTGACCCACAGTGAAGAAACAGATTCCTGCGAGGAGACTCAGGGGGACCCCTC 1206
 Oy 1141 GTCTGTCCCTCCAAAGCCGCAATGACTTGAATGATGAGCTGGGGCCGTGATGT 1200
 Db 1207 GTCTGTCCCTCCAAAGCCGCAATGACTTGAATGATGAGCTGGGGCCGTGATGT 1266
 Oy 1201 GCCCTGAAGAGACAGCAGGCGCTTACAGAGAGTCTCACACTTTCACCTGATCCGC 1260
 Db 1267 GCCCTGAAGAGACAGCAGGCGCTTACAGAGAGTCTCACACTTTCACCTGATCCGC 1326
 Oy 1261 AGTCACCAAGAAAGAGATGCGCTGCGCTTGA 1296
 Db 1327 AGTCACCAAGAAAGAGATGCGCTGCGCTTGA 1362

RESULT 5
 ABQ92264
 ID ABQ92264 standard; cDNA; 2294 BP.
 XX
 AC ABQ92264;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human lung cancer associated cDNA sequence SEQ ID NO:123.
 XX
 KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200247534-A2.
 XX
 PD 20-JUN-2002.
 XX
 PF 30-NOV-2001; 2001WO-US047576.
 XX
 PR 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 PR 28-JUN-2001; 2001US-00897778.
 XX
 PA (CORI-) CORIXA CORP.
 PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vadvick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
 XX
 DR WPI; 2002-583465/62.
 XX
 PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
 PT the polynucleotides, useful in pharmaceutical compositions such as

PR 28-JUN-2000; 2000US-00606421.
 PR 02-AUG-2000; 2000US-00630940.
 PR 21-AUG-2000; 2000US-00643597.
 PR 15-SEP-2000; 2000US-00662786.
 PR 09-OCT-2000; 2000US-00685696.
 PR 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 PR 28-JUN-2001; 2001US-00897778.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI Monnell PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
 XX
 DR WPI; 2003-540796/51.
 XX
 PT New isolated polynucleotides and polypeptides useful for diagnosing,
 PT preventing and/or treating cancer, particularly lung cancer.
 XX
 PS Example 3; Page 132-133; 296pp; English.
 XX
 CC The invention describes isolated polynucleotides and polypeptides useful
 CC for diagnosing, preventing and/or treating cancer, particularly lung
 CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
 CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
 CC specification; complements of the nucleotide sequences cited above; at
 CC least 10 contiguous residues of the nucleotide sequences cited above; a
 CC sequence that hybridize to any of the nucleotide sequences under highly
 CC stringent conditions; a sequence that is at least 75 or 90% identical to
 CC the above nucleotide sequences; or degenerate variants of the above
 CC nucleotide sequences. The composition and methods are useful in
 CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
 CC in gene therapy and in vaccines. This sequence represents a contig cDNA
 CC isolated from human lung tumour cells.
 XX
 SQ Sequence 2294 BP; 549 A; 596 C; 604 G; 545 T; 0 U; 0 Other:
 Query Match 100.0%; Score 1296; DB 8; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 2.9e-267;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGAGCCCTGTGGCGGCTGCTTCTGTGGCTCTGCGTGGAGCGATCCAAAGC 60
 DB 67 ATGAGAGCCCTGTGGCGGCTGCTTCTGTGGCTCTGCGTGGAGCGATCCAAAGC 126
 QY 61 AGCAATGAATTGATCAAGTTCCATCGAATCTGACTGTCTAAATGAGAGCATGTGTG 120
 DB 127 AGCAATGAATTGATCAAGTTCCATCGAATCTGACTGTCTAAATGAGAGCATGTGTG 186
 QY 121 TCCAAAGATGATCTTTCGAATTCACCTGTGCACTGGCCAAAGAAATTCGAGGCGAG 180
 DB 187 TCCAAAGATGATCTTTCGAATTCACCTGTGCACTGGCCAAAGAAATTCGAGGCGAG 246
 QY 181 CACTGTGAATATAGATAGTCAAAACCTGTATGAGGGGAATGTCACCTTTACCGAGGA 240
 DB 247 CACTGTGAATATAGATAGTCAAAACCTGTATGAGGGGAATGTCACCTTTACCGAGGA 306
 QY 241 AAGGCGAGCACTGACACATGAGCGCGGCTGCTGCTGAGCACTGCGACTGTCTT 300
 DB 307 AAGGCGAGCACTGACACATGAGCGCGGCTGCTGCTGAGCACTGCGACTGTCTT 366
 QY 301 CAGCAAACTGATCCATGCGCCACAGATCTGATGCTTTCAAGTGGGCTGGGGAACATTAAT 360
 DB 367 CAGCAAACTGATCCATGCGCCACAGATCTGATGCTTTCAAGTGGGCTGGGGAACATTAAT 426
 QY 361 TACTGAGAGAACCCAGACCAACCGAGGCGACCTGTGATGATGAGGAGGCGCTTAAG 420
 DB 427 TACTGAGAGAACCCAGACCAACCGAGGCGACCTGTGATGATGAGGAGGCGCTTAAG 486
 QY 421 CCGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 487 CCGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546

QY 481 CCAGAGAAATTAATTAATTTAGTGTGGCCAAAGACTGTGAGGCGGCTTTAAGATTATT 540
 DB 547 CCAGAGAAATTAATTAATTTAGTGTGGCCAAAGACTGTGAGGCGGCTTTAAGATTATT 606
 QY 541 GGGGGAGAAATTCACCAACATCGAGAACAGCCCTGTGTTGGGCAATCTACAGAGGCGAC 600
 DB 607 GGGGGAGAAATTCACCAACATCGAGAACAGCCCTGTGTTGGGCAATCTACAGAGGCGAC 666
 QY 601 CCGGGGGGGCTGTGTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 660
 DB 667 CCGGGGGGGCTGTGTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 726
 QY 661 AGGCGCAACACGCTTCATGATTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 720
 DB 727 AGGCGCAACACGCTTCATGATTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 786
 QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTTAGGTTGAGAAACCTTAATC 780
 DB 787 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTTAGGTTGAGAAACCTTAATC 846
 QY 781 CTACACAGAGCTACACAGCGCTGACAGCGCTTGTACCAACAGACATTTGCTTGTGAAG 840
 DB 847 CTACACAGAGCTACACAGCGCTGACAGCGCTTGTACCAACAGACATTTGCTTGTGAAG 906
 QY 841 ATCCGTTCCAAAGAGGCGAGGTGTGGGAGCGCATCCCGCATTAACAGACATCTGCGTG 900
 DB 907 ATCCGTTCCAAAGAGGCGAGGTGTGGGAGCGCATCCCGCATTAACAGACATCTGCGTG 966
 QY 901 CCTCATGATATACATCCATCCAGTTTGGGCAAGCTGTGATGATCACTGGCTTTGGAAAA 960
 DB 967 CCTCATGATATACATCCATCCAGTTTGGGCAAGCTGTGATGATCACTGGCTTTGGAAAA 1026
 QY 961 GAGAAATTTACCGATATCTTATCCGAGACGCTGAAATGACTGTGTGAAGCTGATT 1020
 DB 1027 GAGAAATTTACCGATATCTTATCCGAGACGCTGAAATGACTGTGTGAAGCTGATT 1086
 QY 1021 TCCCAACGGGAGGTGTGAGAGCCGCCATCTACAGGCTGTGAAGTCCACCAAAATGCTG 1080
 DB 1087 TCCCAACGGGAGGTGTGAGAGCCGCCATCTACAGGCTGTGAAGTCCACCAAAATGCTG 1146
 QY 1081 TGTGTGTGTGACCCACAGATGAGAAACAGATTCTGTGACAGGAGACTCAAGGGGAGACCCCTC 1140
 DB 1147 TGTGTGTGTGACCCACAGATGAGAAACAGATTCTGTGACAGGAGACTCAAGGGGAGACCCCTC 1206
 QY 1141 GTCTGTTCCTCCCAAGGCGCATGACTTTGACTGAATTTGTGAGCTGGGCGGTGATGT 1200
 DB 1207 GTCTGTTCCTCCCAAGGCGCATGACTTTGACTGAATTTGTGAGCTGGGCGGTGATGT 1266
 QY 1201 GCCCTGAAGAGCAAGCCAGCGCTTACAGAGAGTCTCACACTTTTACCTTGAATCCGC 1260
 DB 1267 GCCCTGAAGAGCAAGCCAGCGCTTACAGAGAGTCTCACACTTTTACCTTGAATCCGC 1326
 QY 1261 AGTCACCAAGAGAGAGAGATGAGCTGTGCGCTTGA 1296
 DB 1327 AGTCACCAAGAGAGAGATGAGCTGTGCGCTTGA 1362
 RESULT 7
 AD853418
 ID AD853418 standard; cDNA; 2294 BP.
 XX
 AC AD853418;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human lung tumour protein cDNA sequence #120.
 XX
 KM Human; lung tumour protein; lung cancer; cytosolic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2003119763-A1.


```
Key      Location/Qualifiers
FH CDS    70..1365
FT       /*tag= a
FT       /product= "peptide"
FT       130..1365
FT       /*tag= b
FT       misc_feature
FT       2298
FT       /tag= C
FT       /note= "polyA site"
PN       EP288435-A.
XX       26-OCT-1988.
XX
XX       11-APR-1988; 88EP-00810234.
XX
XX       15-APR-1987; 87GB-00009081.
XX       16-JUN-1987; 87GB-00014059.
XX       04-DEC-1987; 87IE-00003299.
XX
PA       (CIBA ) CIBA GEIGY AG.
XX
XX       Meyhack B, Heim J, Burgi R;
XX
XX       WPI: 1988-301440/43.
XX       P-PSDB; AAP80430.
XX
XX       Prod. of human single chain urokinase-type plasminogen activator - by
XX       PT culturing yeast strain transformed with hybrid vector contg. yeast
XX       PT expression control sequences.
XX
XX       Example 1; Fig 2; 48bp; English.
XX
XX       The patent is for the prodn. of human single chain urokinase-type
XX       CC plasminogen activator (UTPA). Mutants of scu-PA are especially those
XX       CC which render the protein protease resistant. The UTPA proteins exhibit
XX       CC the biological activity of natural human UTPA without any refolding
XX       CC procedure being necessary. They can be used as for known PAs in humans
XX       CC for the prevention or treatment of thrombosis or other conditions where
XX       CC it is desired to produce local fibrinolytic or proteolytic activity.
XX       CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ       Sequence 2298 BP; 548 A; 594 C; 607 G; 548 T; 0 U; 1 Other;
Query Match      99.9%; Score 1294.4; DB 1; Length 2298;
Best Local Similarity 99.9%; Pred. No. 6.4e-267;
Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1  ATGAGAGCCCTGCTGGGCGGCGCTTCTCTGCTCTGCTGCTGAGCGAATCCAAAGGC 60
DB      70  ATGAGAGCCCTGCTGGGCGGCGCTTCTCTGCTCTGCTGCTGAGCGAATCCAAAGGC 129
QY      61  AGCAATGAATCTTATCATCAAGTTCCATCGAATGTAATGAGAGAGATGTGTG 120
DB      130  AGCAATGAATCTTATCAAGTTCCATCGAATGTAATGAGAGAGATGTGTG 189
QY      121  TCCAAACAAGTACTTCCAAACATTCAGTGTGCACTGCCAAAGAAATTGAGAGGCGAG 180
DB      190  TCCAAACAAGTACTTCCAAACATTCAGTGTGCACTGCCAAAGAAATTGAGAGGCGAG 249
QY      181  CACTGTGAATATAGTAAGTAAAGAACTGCTATGAGGGGAAATGTCACTTTACCGAGGA 240
DB      250  CACTGTGAATATAGTAAGTAAAGAACTGCTATGAGGGGAAATGTCACTTTACCGAGGA 309
QY      241  AAGGCGAGCACTAGACCAATGAGCGCGGCTGCTGCTGCAATCTGCACTGTCTCTT 300
DB      310  AAGGCGAGCACTAGACCAATGAGCGCGGCTGCTGCTGCAATCTGCACTGTCTCTT 369
QY      301  CAGCAAAACGTACATGCGCCACAGATCTGATGCTCTTCACTGGGCTGGGAGAAACATTAAT 360
DB      370  CAGCAAAACGTACATGCGCCACAGATCTGATGCTCTTCACTGGGCTGGGAGAAACATTAAT 429
QY      361  TACTGACAGAACCAAGCAACCGGAGCGACCTGTGTGCTATGTGACAGGTGGGCTTAAG 420
```

```
DB      430  TACTGACAGAACCAAGCAACCGGAGCGACCTGTGTGCTATGTGACAGGTGGGCTTAAG 489
QY      421  CGCCTTGTCCAAAGATGTCATGTGCTACTGTGCGCAGATGAGAAAAAGCCCTCTCTCT 480
DB      490  CGCCTTGTCCAAAGATGTCATGTGCTACTGTGCGCAGATGAGAAAAAGCCCTCTCTCT 549
QY      481  CCAGAAATTAATTAATTTAGTGTGGCCAAAGACCTGTAGGCGCCGCTTAAGATTAT 540
DB      550  CCAGAAATTAATTAATTTAGTGTGGCCAAAGACCTGTAGGCGCCGCTTAAGATTAT 609
QY      541  GGGGAGAAATTCACCAACATTCAGAAACCAAGCCCTGTGTTGGCGCATCTACAGAGGAC 600
DB      610  GGGGAGAAATTCACCAACATTCAGAAACCAAGCCCTGTGTTGGCGCATCTACAGAGGAC 669
QY      601  CGGGGGGGGCTCTGTCACTTACGTGTGTGAGAGGAGCCCTCATACCCCTTGTGGGTATC 660
DB      670  CGGGGGGGGCTCTGTCACTTACGTGTGTGAGAGGAGCCCTCATACCCCTTGTGGGTATC 729
QY      661  AGGCGCACACACTGCTTCAATGATTAACCAAGAGAGGACATACCTGCTACCTGGGT 720
DB      730  AGGCGCACACACTGCTTCAATGATTAACCAAGAGAGGACATACCTGCTACCTGGGT 789
QY      721  CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTTGAGGTGAGAAAACCTATC 780
DB      790  CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTTGAGGTGAGAAAACCTATC 849
QY      781  CTACACAGAGCTACAGCGCTGACACGCTTGTCTACACACAGACATTCGCTTGTGAAG 840
DB      850  CTACACAGAGCTACAGCGCTGACACGCTTGTCTACACACAGACATTCGCTTGTGAAG 909
QY      841  ATCCGTTCCAAAGAGGAGGAGGTGTGGCAGCCATCCCGGACTTACAGACATCTGCTG 900
DB      910  ATCCGTTCCAAAGAGGAGGAGGTGTGGCAGCCATCCCGGACTTACAGACATCTGCTG 969
QY      901  CCTCGATGTATACGATCCCCAGTTTGGCAACAAGCTGTGATCACTGGCTTTGGAGAAA 960
DB      970  CCTCGATGTATACGATCCCCAGTTTGGCAACAAGCTGTGATCACTGGCTTTGGAGAAA 1029
QY      961  GAGAAATTTACCGACTATCTCTATCCGAGACAGCTGAATAATGACTGTTGAACTGAT 1020
DB      1030  GAGAAATTTACCGACTATCTCTATCCGAGACAGCTGAATAATGACTGTTGAACTGAT 1089
QY      1021  TCCCAACGGAGAGGTGAGAGCCCACTACTACGGCTCTGAATCAACCAAAATGCTG 1080
DB      1090  TCCCAACGGAGAGGTGAGAGCCCACTACTACTGAGCTCTGAATCAACCAAAATGCTG 1149
QY      1081  TGTGCTGTGACCCACAGTGAATAACAGATTCTGTCCAGGGAGATCAAGGGGAGACCCCTG 1140
DB      1150  TGTGCTGTGACCCACAGTGAATAACAGATTCTGTCCAGGGAGATCAAGGGGAGACCCCTG 1209
QY      1141  GTCTGTTCCTCCAGAGCGCGCATGACTTGTGACTGAATTTGTAGTGTGGGCGCTGTAGAT 1200
DB      1210  GTCTGTTCCTCCAGAGCGCGCATGACTTGTGACTGAATTTGTAGTGTGGGCGCTGTAGAT 1269
QY      1201  GCCCTGAAGAGCAAGCCAGGCGCTTACAGAGAGTCTACACATTTTACCTTGGATCCGC 1260
DB      1270  GCCCTGAAGAGCAAGCCAGGCGCTTACAGAGAGTCTACACATTTTACCTTGGATCCGC 1329
QY      1261  AGTCACCAAGGAAGAGATGAGCTGTGGCCCTTGA 1296
DB      1330  AGTCACCAAGGAAGAGATGAGCTGTGGCCCTTGA 1365
RESULT 9
AA006049
ID      AA006049 standard; DNA; 1296 BP.
XX
XX      AA006049;
AC      25-MAR-2003 (revised)
DT      24-JAN-1991 (first entry)
XX
```

DE plasmid pUK1 pro-Urokinase sequence.
XX
KW pro-Urokinase; transgenic mice; ss.
XX
OS Synthetic.
XX
PN EP30592-A.
XX
PD 03-OCT-1990.
XX
PF 30-MAR-1990; 90EP-00303445.
XX
PR 31-MAR-1989; 89JP-00078574.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (EXPE-) CENT INST EXPER ANIMALS.
PA (JIKK-) JIKKEN DOBUTSU CHOU KENK.
XX
PI Sekine S, Ito S, Katsuki M;
XX
DR WPI; 1990-299492/40.
DR P-PSDB: AAR07112.
XX
PT Prodn. of recombinant protein, esp. human pro:urokinase - from milk of
PT transgenic animals using promoter of bovine alpha S1 casein chromosomal
PT gene.
XX
PS Example; Table 1; 55pp; English.
XX
CC E.coli strain C600SF8 was transformed with recombinant plasmid containing
CC the DNA derived from human pharynx cancer cell strain Detroit 562. 10000
CC colonies were screened with the probe and one positive clone was
CC identified. Plasmid pUK1 was isolated and found to contain the coding
CC region and 3' non-coding region of pro-UK downstream of Cys(44). Four
CC silent substitutions were identified c.f. Holmes et al., Biotechnology,
CC vol.3, p.923 (1985) as follows: Asn(254), AAC to AAT; Leu(340), CTA to
CC CTG; Pro(345), CCC to CCA; Gln(346), CAA to CAG. See also AAO06045-006048
CC and AAO06392. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1296 BP; 327 A; 361 C; 337 G; 271 T; 0 U; 0 Other;

Query Match 99.8%; Score 1292.8; DB 2; Length 1296;
Best Local Similarity 99.8%; Pred. No. 1.4e-266;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGCGCCCTGCTCTCTGCGCTCTGTCGTAAGCACTCCAAAGG 60
DB 1 ATGAGAGCCCTGCTGCGCGCCCTGCTCTCTGCGCTCTGTCGTAAGCACTCCAAAGG 60

QY 61 AGCAATGACTTTCATCAAGTTCCATGCACTGACTGTCTAAATGAGAGACATGTGTG 120
DB 61 AGCAATGACTTTCATCAAGTTCCATGCACTGACTGTCTAAATGAGAGACATGTGTG 120

QY 121 TCCAAAGAGTACTTTCCTCAACATTCCTGCTGCAATCGCCCAAGAAATTCGAGGGCAG 180
DB 121 TCCAAAGAGTACTTTCCTCAACATTCCTGCTGCAATCGCCCAAGAAATTCGAGGGCAG 180

QY 181 CACTGTGAATAGATAAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTTACGAGGA 240
DB 181 CACTGTGAATAGATAAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTTACGAGGA 240

QY 241 AAGGCCGACCTGACACATGAGGCGCGCCCTGCTGCGAACTCTGCACTGTCTT 300
DB 241 AAGGCCGACCTGACACATGAGGCGCGCCCTGCTGCGAACTCTGCACTGTCTT 300

QY 301 CAGCAAACTGACATGCGCCAGATCTGATCTCTTCAAGTGGGCTTGAGGAAACATTAAT 360
DB 301 CAGCAAACTGACATGCGCCAGATCTGATCTCTTCAAGTGGGCTTGAGGAAACATTAAT 360

QY 361 TACTGACGAAACCCAGCAACCGAGAGCGAACCCTGTGTCTATGTGAGGTGGGCTTAAG 420
DB 361 TACTGACGAAACCCAGCAACCGAGAGCGAACCCTGTGTCTATGTGAGGTGGGCTTAAG 420

QY 421 CCGCTTGTCCAAAGATGCATGCTGCATGACTCGCAGATGAAAAAGCCCTCTCTCT 480
DB 421 CCGCTTGTCCAAAGATGCATGCTGCATGACTCGCAGATGAAAAAGCCCTCTCTCT 480

QY 481 CCAGAGAAATTTAAATTTAGTGTGGCCAAAAGACTCTGAGGGCCCGCTTTAAGATTAT 540
DB 481 CCAGAGAAATTTAAATTTAGTGTGGCCAAAAGACTCTGAGGGCCCGCTTTAAGATTAT 540

QY 541 GGGGGGAAATTCACCAACATCGAGAACAGCCCTGTTTGGCGCCATCTACAGAGGAC 600
DB 541 GGGGGGAAATTCACCAACATCGAGAACAGCCCTGTTTGGCGCCATCTACAGAGGAC 600

QY 601 CGGGGGGGCTCTGTCACTACCTGTGTGTGAGGAGGACCTCTATAGCCCTTGTGGTATC 660
DB 601 CGGGGGGGCTCTGTCACTACCTGTGTGTGAGGAGGACCTCTATAGCCCTTGTGGTATC 660

QY 661 AGCGCCACACTGCTCTTATGATTAACCAAGAGAGAGACTATACGTCTACCTGAGT 720
DB 661 AGCGCCACACTGCTCTTATGATTAACCAAGAGAGAGACTATACGTCTACCTGAGT 720

QY 721 CGCTCAAGCTTAACTCCACACGCGAAGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 780
DB 721 CGCTCAAGCTTAACTCCACACGCGAAGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 780

QY 781 CTACACAAAGGACTACAGGCTGACACGCTGTGTCCACCAACGACATTCCTGCTGAG 840
DB 781 CTACACAAAGGACTACAGGCTGACACGCTGTGTCCACCAACGACATTCCTGCTGAG 840

QY 841 ATCCGTTCCAAAGAGAGGAGGAGGTGTGCGAGCCATCCCGGACTAATACAGACCATGCTG 900
DB 841 ATCCGTTCCAAAGAGAGGAGGAGGTGTGCGAGCCATCCCGGACTAATACAGACCATGCTG 900

QY 901 CCTTGATGTATTAAGATATCCCAAGTTTGGACCAAGCTGTGATCATGCTTTGGAATA 960
DB 901 CCTTGATGTATTAAGATATCCCAAGTTTGGACCAAGCTGTGATCATGCTTTGGAATA 960

QY 961 GAGAAATTCACGCACTATCTCTATCCGAGAGAGCTGAAAAAGACTGTTGTGAAGTGAAT 1020
DB 961 GAGAAATTCACGCACTATCTCTATCCGAGAGAGCTGAAAAAGACTGTTGTGAAGTGAAT 1020

QY 1021 TCCCAACGGGAGTGTACAGACCCCACTACTAGGCTCTGAAATGACCAACCAAAATGCTG 1080
DB 1021 TCCCAACGGGAGTGTACAGACCCCACTACTAGGCTCTGAAATGACCAACCAAAATGCTG 1080

QY 1081 TGTGCTGTGACCCCAAGTGTGAATAAGATTCTGCGACAGGAGACTAGAGGGGACCCCTC 1140
DB 1081 TGTGCTGTGACCCCAAGTGTGAATAAGATTCTGCGACAGGAGACTAGAGGGGACCCCTC 1140

QY 1141 GTCTGTTCCTCCAAAGGCGCGATGACTTTGACTGGAATTGTGAGCTGGGCGCTGATGT 1200
DB 1141 GTCTGTTCCTCCAAAGGCGCGATGACTTTGACTGGAATTGTGAGCTGGGCGCTGATGT 1200

QY 1201 GCCCTGAAGGACCAAGCGGCTGTACACGAGATCTCAACTTTTACCTTGATTCGC 1260
DB 1201 GCCCTGAAGGACCAAGCGGCTGTACACGAGATCTCAACTTTTACCTTGATTCGC 1260

QY 1261 AGTCACCAAGAGAGAAATGAGCTGAGCCCTCTGA 1296
DB 1261 AGTCACCAAGAGAGAAATGAGCTGAGCCCTCTGA 1296

RESULT 10
AA055772
ID AA055772 standard; cDNA to mRNA; 1296 BP.
XX
XX AA055772;
AC
XX 13-JUN-1994 (first entry)
DT
XX Pro-urokinase derivative.
DE
XX Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor;
KW 89.

XX 06-OCT-1989.
 PD
 PF 30-MAR-1988; 88JP-00079738.
 XX
 PR 30-MAR-1988; 88JP-00079738.
 XX
 PA (GREC) GREEN CROSS CORP.
 XX
 DR WPI; 1989-337032/46.
 DR P-PSDB; AAP93256.
 XX
 PT Recovery of variant human prourokinase (PUK) - comprises solubilisation
 PT treatment using guanidine hydrochloride of ppte. obd. from centrifuged
 PT prod. from transformed yeast culture.
 PS
 XX Disclosure; Fig 3(1)-(2); 11pp; Japanese.
 CC Variant human PUK is expressed from plasmid pUH33 in *S.cerevisiae* AH22.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 1473 BP; 357 A; 429 C; 385 G; 302 T; 0 U; 0 Other;

Query Match 99.8%; Score 1292.8; DB 1; Length 1473;
 Best Local Similarity 99.8%; Pred. No. 1.4e-266;
 Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGGCGGCGCCCTGCTTCTCTGGCTCTGCTGCTGAGCCGATCCAAAGC 60
 DB 81 ATAGAGCCCTGCTGGCGGCGCCCTGCTTCTGCGTCTGCTGCTGAGCCGATCCAAAGC 140
 QY 61 AGCAATGAATCTTCATCAAGTTCATGCACTGAGTGTCTTAATGAGAGAAATGATG 120
 DB 141 AGCAATGAATCTTCATCAAGTTCATGCACTGAGTGTCTTAATGAGAGAAATGATG 200
 QY 121 TCCAACAAGTACTTCTTCCAACTTCACTGCTGCACTGCTGCACTGCTGCACTGCTGCA 180
 DB 201 TCCAACAAGTACTTCTTCCAACTTCACTGCTGCACTGCTGCACTGCTGCACTGCTGCA 260
 QY 181 CACTGGAATTAATAGTAAAGTAAACCTGTAAGAGAGAAATGTAATTTACGAGAG 240
 DB 261 CACTGGAATTAATAGTAAAGTAAACCTGTAAGAGAGAAATGTAATTTACGAGAG 320
 QY 241 AAGGCCAGCATGACATGAGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 321 AAGGCCAGCATGACATGAGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
 QY 301 CAGCAACGTAACATGCGCCAGATGATGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 381 CAGCAACGTAACATGCGCCAGATGATGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 440
 QY 361 TACTGAGAGAACCCAGCAACCGGAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 441 TACTGAGAGAACCCAGCAACCGGAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500
 QY 421 CCGCTTGTCCAAAGTGCATGCTGATGACTGCGCAGATGAGAAAGAGCCCTCTCTCT 480
 DB 501 CCGCTTGTCCAAAGTGCATGCTGATGACTGCGCAGATGAGAAAGAGCCCTCTCTCT 560
 QY 481 CCGAAGAAATTAATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 561 CCGAAGAAATTAATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620
 QY 541 GGGGAGAAATTCACCACTGAGAACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 DB 621 GGGGAGAAATTCACCACTGAGAACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
 QY 601 CCGGGGGGCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 DB 681 CCGGGGGGCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740
 QY 661 AAGGCACACACTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

DB 741 AAGGCACACACTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
 QY 721 CGCTCAAGGCTTAATTCACCACTGAGAACCGGAGATGAGATTGAGTGGAAAACTTAATC 780
 DB 801 CGCTCAAGGCTTAATTCACCACTGAGAACCGGAGATGAGATTGAGTGGAAAACTTAATC 860
 QY 781 CTACACAGAGACTACAGGCTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 861 CTACACAGAGACTACAGGCTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 QY 841 ATCCGTTCCAAAGAGAGGAGTGTGCGAGCATCCGAGATTAACAGACCATCTGCTGCT 900
 DB 921 ATCCGTTCCAAAGAGAGGAGTGTGCGAGCATCCGAGATTAACAGACCATCTGCTGCT 980
 QY 901 CCTTGATGATTAACGATCCCAAGTTTGGCAAGAGCTGTGATCACTGCTTGGAAAA 960
 DB 981 CCTTGATGATTAACGATCCCAAGTTTGGCAAGAGCTGTGATCACTGCTTGGAAAA 1040
 QY 961 GAGAAATTCACCGACTATCTCTATCCGAGAGAGCTGMAAATGACTGTTGAGCTGATT 1020
 DB 1041 GAGAAATTCACCGACTATCTCTATCCGAGAGAGCTGMAAATGACTGTTGAGCTGATT 1100
 QY 1021 TCCCAACCGAGAGTGTGACAGACCCCACTACTACTGAGCTGTGAAGTACCAACCAATGCTG 1080
 DB 1101 TCCCAACCGAGAGTGTGACAGACCCCACTACTACTGAGCTGTGAAGTACCAACCAATGCTG 1160
 QY 1081 TGTGCTGTGACCCCACTGAGTGAAGAAACAGATTCTGCTGCAAGGAGACTGAGGGGAGCCCTTC 1140
 DB 1161 TGTGCTGTGACCCCACTGAGTGAAGAAACAGATTCTGCTGCAAGGAGACTGAGGGGAGCCCTTC 1220
 QY 1141 GTCTGTTCCCTCAAGGCGCGCATGACTTGTGAATTTGAGCTGAGGGGCGGTGAGATGT 1200
 DB 1221 GTCTGTTCCCTCAAGGCGCGCATGACTTGTGAATTTGAGCTGAGGGGCGGTGAGATGT 1280
 QY 1201 GCCCTGAAGAGCAAGCCAGGCGTCTACACGAGAGTCTCACATTTCTTACCTGTGATCCGC 1260
 DB 1281 GCCCTGAAGAGCAAGCCAGGCGTCTACACGAGAGTCTCACATTTCTTACCTGTGATCCGC 1340
 QY 1261 AGTCACACCAAGAGAGAAATGCGCTGCGCTCTCTGA 1296
 DB 1341 AGTCACACCAAGAGAGAAATGCGCTGCGCTCTCTGA 1376

RESULT 12
 AAN50138
 ID AAN50138 standard; DNA; 1474 BP.
 XX
 AC AAN50138;
 XX
 DT 27-SEP-1991 (first entry)
 XX
 DE Sequence of the signal sequence and noncoding region of the pro-UK
 DE structural gene (Sequence II).
 XX
 KW Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT sig_peptide 81..140
 FT /*tag= a
 FT mat_peptide 141..1376
 FT /*tag= b
 FT 141..611
 FT /*tag= c
 FT /product= "A chain"
 FT CDS 615..1376
 FT /*tag= d
 FT /product= "B chain"
 XX
 XX EPI54272-A.
 XX

XX Read SG, Wang T;

XX WPI; 1999-571839/48.

PT New isolated lung tumor polynucleotides, used to develop products for the
PT treatment, prevention and monitoring the progression of lung cancer.

PS Claim 12; Page 114-115; 148pp; English.

XX The invention provides isolated human lung tumor nucleic acids and
CC polypeptides. The polypeptides can be used for the treatment of lung
CC cancer. The polypeptides and polynucleotides can be used to stimulate T
CC cells or antigen presenting cells for use in the treatment of lung
CC cancer. The polypeptides and monoclonal antibodies specific for the
CC polypeptides can also be used to inhibit the development of lung cancer.
CC Agents which bind the polypeptides can be used for detecting lung cancer
CC and for monitoring the progression of lung cancer

XX Sequence 1475 BP; 357 A; 429 C; 386 G; 303 T; 0 U; 0 Other;

Query Match 99.8%; Score 1292.8; DB 2; Length 1475;

Beet Local Similarity 99.8%; Pred. No. 1.4e-266;

Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGAGCCCTGCTGGGCGCGCTCTCTCTGCGGCTCCTGGTGGGAGCGACTCCAAAGC 60
DB 81 ATGAGAGCCCTGCTGGGCGCGCTCTCTCTGCGGCTCCTGGTGGGAGCGACTCCAAAGC 140
QY 61 AGCAATGAATCTTCATCAAGTTCCATGCACTGTGACTGTCTAAATGGAGGACATGTGT 120
DB 141 AGCAATGAATCTTCATCAAGTTCCATGCACTGTGACTGTCTAAATGGAGGACATGTGT 200
QY 121 TCCAAAGAACTCTTCCCAATTCATCTGTGTGCACTGCCCAAGAAATTCGAGGCGAG 180
DB 201 TCCAAAGAACTCTTCCCAATTCATCTGTGTGCACTGCCCAAGAAATTCGAGGCGAG 260
QY 181 CACTGTGAATAGTAAGTCAAAAACTGCTATGAGGGAATGCTACTTTTACGAGGA 240
DB 261 CACTGTGAATAGTAAGTCAAAAACTGCTATGAGGGAATGCTACTTTTACGAGGA 320
QY 241 AAGGCCGACATGACACCATGAGCGCGGCTCTGCTGCTGCAACTCTGCACTGTCTT 300
DB 321 AAGGCCGACATGACACCATGAGCGCGGCTCTGCTGCTGCAACTCTGCACTGTCTT 380
QY 301 CAGCAAACTATACATGCGCCACAGATCTGATGCTCTTCACTGGGCTTGGGAAACATAAT 360
DB 381 CAGCAAACTATACATGCGCCACAGATCTGATGCTCTTCACTGGGCTTGGGAAACATAAT 440
QY 361 TACTGCAAGAACCCAGAACCCGAGGCGACCTGTGTCTATGTGAGGTGGCTTAAG 420
DB 441 TACTGCAAGAACCCAGAACCCGAGGCGACCTGTGTCTATGTGAGGTGGCTTAAG 500
QY 421 CCGCTTGTCCAAAGTGCATGTGCTGCTGCTGCGCAGATGAAAGCCCTCTCTCT 480
DB 501 CCGCTTGTCCAAAGTGCATGTGCTGCTGCTGCGCAGATGAAAGCCCTCTCTCTCT 560
QY 481 CCAGAAAGATTAAATTTAGTGTGCGCAAAAGCTCTGAGGCGCGCTTTAAGATTATT 540
DB 561 CCAGAAAGATTAAATTTAGTGTGCGCAAAAGCTCTGAGGCGCGCTTTAAGATTATT 620
QY 541 GGGGGAGAAATTCACACCATCGAAGAACAGCCCTGTGTTGGCGCATCTACAGAGGAC 600
DB 621 GGGGGAGAAATTCACACCATCGAAGAACAGCCCTGTGTTGGCGCATCTACAGAGGAC 680
QY 601 CGGGGGGGCTCTGTCACTACGTGTGTGAGGAGAGGCTCATCAGCCCTTGTGTGTATC 660
DB 681 CGGGGGGGCTCTGTCACTACGTGTGTGAGGAGAGGCTCATCAGCCCTTGTGTGTATC 740
QY 661 AGGCGCACACACTGCTTCAATTGATACCCAAAGAGAGAGACTACATGCTACCTGGGT 720
DB 741 AGGCGCACACACTGCTTCAATTGATACCCAAAGAGAGAGACTACATGCTACCTGGGT 800

```

```

QY 721 CGCTCAGGCTTAATCTCCAAACGCAAGGGAGATGAAGTTTGAAGTGGAAAACTTAATC 780
DB 801 CGCTCAGGCTTAATCTCCAAACGCAAGGGAGATGAAGTTTGAAGTGGAAAACTTAATC 860
QY 781 CTACACAGGAGCTTACAGAGGCTGACACGCTGTGTACCCAAACGACATGCTGTGTAG 840
DB 861 CTACACAGGAGCTTACAGAGGCTGACACGCTGTGTACCCAAACGACATGCTGTGTAG 920
QY 841 ATCCGTTCCAGAGAGGAGAGGTGTGCGAGCCATCCCGAATATACAGACCATCTGCCTG 900
DB 921 ATCCGTTCCAGAGAGGAGAGGTGTGCGAGCCATCCCGAATATACAGACCATCTGCCTG 980
QY 901 CCTCGATGATTAAGATATCCCAAGTTTGGCACAACTGTGATGATCTGCTTGGAAAA 960
DB 981 CCTCGATGATTAAGATATCCCAAGTTTGGCACAACTGTGATGATCTGCTTGGAAAA 1040
QY 961 GAGAAATTCACAGCACTATCTGATCCGAGAGAGCTGAAAAATGACTGTTGTGAGCTGATT 1020
DB 1041 GAGAAATTCACAGCACTATCTGATCCGAGAGAGCTGAAAAATGACTGTTGTGAGCTGATT 1100
QY 1021 TCCCAACGGAGAGTGTACAGAGCCCACTACTACGCTCTGAAATGCAACCAAAATGCTG 1080
DB 1101 TCCCAACGGAGAGTGTACAGAGCCCACTACTACGCTCTGAAATGCAACCAAAATGCTG 1160
QY 1081 TGTGCTGTGACCCCAAGTGGAAAAAGATTTCTGCGGAGAGACTCAGGGGAGCCCTTC 1140
DB 1161 TGTGCTGTGACCCCAAGTGGAAAAAGATTTCTGCGGAGAGACTCAGGGGAGCCCTTC 1220
QY 1141 GTCTGTCTCTCCAAAGCGCGCATGACTTTGACTGAAATTTGACTGAGGCGGTGATGT 1200
DB 1221 GTCTGTCTCTCCAAAGCGCGCATGACTTTGACTGAAATTTGACTGAGGCGGTGATGT 1280
QY 1201 GCCCTGAAGACACAGCCAGGCGTCTACAGAGTCTTACACTTTTACCTGTGATCCG 1260
DB 1281 GCCCTGAAGACACAGCCAGGCGTCTACAGAGTCTTACACTTTTACCTGTGATCCG 1340
QY 1261 AGTCACACCAAGAAAGAAATGGCGTGGCCCTCTGGA 1296
DB 1341 AGTCACACCAAGAAAGAAATGGCGTGGCCCTCTGGA 1376

```

RESULT 14

ID AAC65858 standard; cDNA; 1475 BP.

XX AAC65858;

DT 21-FEB-2001 (first entry)

DE Human lung cancer-associated cDNA for config 10.

KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW vaccine; detection; ss.

OS Homo sapiens.

PN MO200061612-A2.

PD 19-OCT-2000.

PF 03-APR-2000; 2000WO-US008896.

PR 02-APR-1999; 99US-00285479.

PR 17-DEC-1999; 99US-00466396.

PR 30-DEC-1999; 99US-00476496.

PR 10-JAN-2000; 2000US-00480884.

PR 22-FEB-2000; 2000US-00510376.

PA (CORI-) CORIXA CORP.

PI Wang T, Fan L;

XX WPI; 2000-628399/60.

PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer.
 XX
 PS Claim 3; Page 382-383; 850pp; English.

XX The invention relates to a gene expression profile comprising one or more
 CC genes (AB234869-AB235692) and generated from a cell type. The cell type
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
 CC endothelium, myometrium microvascular endothelium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 CC for determining the level of RNA expression for a sample, determining the
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 CC expression profile is useful in identifying disease pathologies involving
 CC alterations of gene expression. The assessment of expression profiles may
 CC provide meaningful information with respect to tumour type and stage,
 CC treatment methods, and prognosis. The gene or protein expression profile
 CC may also be used for creating microarrays. The microarray is useful for
 CC genetic and physical mapping of genomes, DNA sequencing, genetic or
 CC medical diagnosis, genotyping of organisms, confirming cell or tissue
 CC identifications and in identifying promising antibiotics, antiviral or
 CC antifungal agents.

SO Sequence 1475 BP; 357 A; 429 C; 386 G; 303 T; 0 U; 0 Other;

Query Match 99.8%; Score 1292.8; DB 6; Length 1475;

Best Local Similarity 99.8%; Pred. No. 1.4e-266;

Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGGCGCTGCTTCTGCGTCTGCTGAGCGATCCAAAGC 60
 DB 81 ATGAGAGCCCTGCTGCGGCGCTGCTTCTGCGTCTGCTGAGCGATCCAAAGC 140
 QY AGCAATGAATTCATCAAGTTCATGCACTGATGCTGTAAATGAGAGCACTGTG 120
 DB 141 AGCAATGAATTCATCAAGTTCATGCACTGATGCTGTAAATGAGAGCACTGTG 200
 QY 121 TCCAAACAAGTACTTCTCAACATTCAGTGTGCACTGCCAAGAAATTCGAGGGCAG 180
 DB 201 TCCAAACAAGTACTTCTCAACATTCAGTGTGCACTGCCAAGAAATTCGAGGGCAG 260
 QY 181 CACTGTGAATTAAGTAAGTAAAACTGCTATGAGGGAGATGTCATTCTTAACGAGGA 240
 DB 261 CACTGTGAATTAAGTAAGTAAAACTGCTATGAGGGAGATGTCATTCTTAACGAGGA 320
 QY 241 AAGGCCAGCACTGACACCATGAGGCGGCGCTGCTGCTGCAACCTGCTGCTCT 300
 DB 321 AAGGCCAGCACTGACACCATGAGGCGGCGCTGCTGCTGCAACCTGCTGCTCT 380
 QY 301 CAGCAAACTGACATGCGCCACAGATCTGATGCTTCACTGAGGCTCTGGGAAACATAAT 360
 DB 381 CAGCAAACTGACATGCGCCACAGATCTGATGCTTCACTGAGGCTCTGGGAAACATAAT 440
 QY 361 TACTGAGAGAACCCAGAACAAACCGAGAGGCACTGCTGCTATGCTGAGGTGGCTTAAG 420
 DB 441 TACTGAGAGAACCCAGAACAAACCGAGAGGCACTGCTGCTATGCTGAGGTGGCTTAAG 500
 QY 421 CCGCTTGTCCAAAGAGTGCATGCTGATGCTGCGAGATGAGAAAGCCCTCTCTCT 480
 DB 501 CCGCTTGTCCAAAGAGTGCATGCTGATGCTGCGAGATGAGAAAGCCCTCTCTCTCT 560
 QY 481 CCGAAGAAATTAATTTCACTGTGCGCAAAAGACTCTGAGGCCCCGCTTTAAGATTAT 540
 DB 561 CCGAAGAAATTAATTTCACTGTGCGCAAAAGACTCTGAGGCCCCGCTTTAAGATTAT 620
 QY 541 GGGGAGAAATTCACACATGAGAACGAGCCCTGCTTGGGGCACTACAGAGAGCAC 600

DB 621 GGGGAGAAATTCACACATGAGAACGAGCCCTGCTTGGGGCACTACAGAGAGCAC 680
 QY 601 CCGGGGGGCTGTGTCACCTAAGTGTGAGAGGACCTTCATCAGCCCTTGTGGGTATC 660
 DB 681 CCGGGGGGCTGTGTCACCTAAGTGTGAGAGGACCTTCATCAGCCCTTGTGGGTATC 740
 QY 661 AGCGGCACACATGCTTCAATGATTAACCAAGAGAGAGACTACATGCTTACCTGGGT 720
 DB 741 AGCGGCACACATGCTTCAATGATTAACCAAGAGAGAGACTACATGCTTACCTGGGT 800
 QY 721 CCGTCAAGGCTTAATTCACACAGGAGGAGATGAAGTTGAGTGGAGAAACCTAATC 780
 DB 801 CCGTCAAGGCTTAATTCACACAGGAGGAGATGAAGTTGAGTGGAGAAACCTAATC 860
 QY 781 CTACACAAGGACTACAGGCTGACACGCTTGTCTACCAACAGCATTCCTGCTGAG 840
 DB 861 CTACACAAGGACTACAGGCTGACACGCTTGTCTACCAACAGCATTCCTGCTGAG 920
 QY 841 ATCGCTTCAGAGAGGAGGAGTGTGCGAGCCATCCGAGCTATACAGACCATCTGCTG 900
 DB 921 ATCGCTTCAGAGAGGAGGAGTGTGCGAGCCATCCGAGCTATACAGACCATCTGCTG 980
 QY 901 CCGCTGATGTAATTAAGATCCCAAGTTTGGCAAGCTGTGATCACTGGCTTTGAAAA 960
 DB 981 CCGCTGATGTAATTAAGATCCCAAGTTTGGCAAGCTGTGATCACTGGCTTTGAAAA 1040
 QY 961 GAGAAATTCACGACTATCTATCCGAGAGAGCTGAAAAATGACTGTGTAAGCTGATT 1020
 DB 1041 GAGAAATTCACGACTATCTATCCGAGAGAGCTGAAAAATGACTGTGTAAGCTGATT 1100
 QY 1021 TCCACACGGAGAGTGTGAGAGCCCACTACTACGAGCTGTGAGTCAACCAAAATGCTG 1080
 DB 1101 TCCACACGGAGAGTGTGAGAGCCCACTACTACGAGCTGTGAGTCAACCAAAATGCTG 1160
 QY 1081 TGTGCTGTGACCAAGTGTGAAAAAGATTCCGTGCAAGGAGACTGAGGGGGACCCCTC 1140
 DB 1161 TGTGCTGTGACCAAGTGTGAAAAAGATTCCGTGCAAGGAGACTGAGGGGGACCCCTC 1220
 QY 1141 GTCTGTTCCTCCAAAGCCGATGACTTTGATGGAATGTGAGCTGGGCGGTGATGT 1200
 DB 1221 GTCTGTTCCTCCAAAGCCGATGACTTTGATGGAATGTGAGCTGGGCGGTGATGT 1280
 QY 1201 GCCCTGAAGGACAGCCAGGCTGTACAGAGAGTCTCAACTTTTACCTTGGATCCGC 1260
 DB 1281 GCCCTGAAGGACAGCCAGGCTGTACAGAGAGTCTCAACTTTTACCTTGGATCCGC 1340
 QY 1261 AGTCAACCAAGAGAAATGAGCTGGCCCTCTGA 1296
 DB 1341 AGTCAACCAAGAGAAATGAGCTGGCCCTCTGA 1376

Search completed: March 21, 2004, 18:17:07
 Job time : 610 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 17:58:59 ; Search time 120 Seconds
(without alignments)
593.470 Million cell updates/sec

Title: US-10-076-421-1

Perfect score: 1296

Sequence: 1 atgagagagccctgctgagcg.....agatgagcctgagcctctga 1296

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfillseq1.seq:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1296	100.0	2294	US-09-643-597-123	Sequence 123, App
2	1296	100.0	2294	US-09-480-884A-123	Sequence 123, App
3	1296	100.0	2294	US-09-542-615A-123	Sequence 123, App
4	1296	100.0	2294	US-09-606-421B-123	Sequence 123, App
5	1296	100.0	2294	US-09-023-655-1217	Sequence 121, App
6	1296	100.0	2294	US-09-221-107-123	Sequence 122, App
7	1292.8	99.8	1475	US-09-643-597-122	Sequence 122, App
8	1292.8	99.8	1475	US-09-480-884A-122	Sequence 122, App
9	1292.8	99.8	1475	US-09-542-615A-122	Sequence 122, App
10	1292.8	99.8	1475	US-09-606-421B-122	Sequence 122, App
11	1292.8	99.8	1475	US-09-023-655-122	Sequence 122, App
12	1292.8	99.8	2301	US-09-221-107-122	Sequence 122, App
13	1288	99.4	1372	5188829-2	Patent No. 5188829
14	1232.8	95.1	1236	US-07-957-039A-7	Sequence 7, Appl
15	1232.8	95.1	1236	US-09-023-655-927	Sequence 927, App
16	1229.6	94.9	1236	US-08-153-799-17	Sequence 17, Appl
17	1223.4	94.4	1233	US-08-254-922-1	Sequence 1, Appl
18	1223.4	94.4	1233	US-08-286-748B-1	Sequence 1, Appl
19	985.8	76.1	1335	US-07-942-157A-2	Sequence 2, Appl
20	600	46.3	600	US-09-101-272G-72	Sequence 72, Appl
21	437.8	33.8	666	US-09-101-272G-97	Sequence 97, Appl
22	405.6	31.3	423	US-08-797-689-11	Sequence 11, Appl
23	405.6	31.3	423	US-09-984-186-11	Sequence 11, Appl
24	396.8	30.6	624	US-09-101-272G-79	Sequence 79, Appl
25	396.8	30.6	645	US-09-101-272G-95	Sequence 95, Appl
26	180.4	13.9	1068	US-08-811-949-58	Sequence 58, Appl
27	169.8	13.1	2033	US-08-148-910-14	Sequence 14, Appl

28	169.8	13.1	2033	1	US-08-448-937A-14	Sequence 14, Appl
29	169.2	13.1	1068	2	US-08-811-949-46	Sequence 46, Appl
30	169.2	13.1	1068	2	US-08-811-949-52	Sequence 52, Appl
31	169.2	13.1	1314	2	US-08-811-949-50	Sequence 50, Appl
32	169.2	13.1	1314	2	US-08-811-949-54	Sequence 54, Appl
33	169.2	13.1	1314	2	US-08-811-949-56	Sequence 56, Appl
34	169.2	13.1	1314	2	US-08-811-949-38	Sequence 38, Appl
35	168.8	13.0	1170	2	US-08-811-949-66	Sequence 66, Appl
36	167.6	12.9	1065	1	US-08-427-640-1	Sequence 1, Appl
37	167.6	12.9	1068	1	US-08-137-116-2	Sequence 2, Appl
38	167.6	12.9	1068	2	US-08-811-949-44	Sequence 44, Appl
39	167.6	12.9	1137	4	US-09-553-498-9	Sequence 9, Appl
40	167.6	12.9	1137	4	US-09-618-869-9	Sequence 9, Appl
41	167.6	12.9	1419	2	US-08-811-949-62	Sequence 62, Appl
42	167.6	12.9	1738	6	5200340-1	Patent No. 5200340
43	167.6	12.9	1848	3	US-08-814-412-10	Sequence 10, Appl
44	167.6	12.9	1955	2	US-08-883-795A-39	Sequence 39, Appl
45	167.6	12.9	2101	2	US-08-811-949-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-09-643-597-123
; Sequence 123, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-123
Query Match 100.0%; Score 1296; DB 4; Length 2294;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAGCCCTGCTGCGCGCTCTCTCTGCTCTGCTGAGCACTCCAAAGC 60
DB ATGAGAGCCCTGCTGCGCGCTCTCTCTGCTCTGCTGAGCACTCCAAAGC 126
QY 61 AGCAATGAATTCATCAATTCATGAGCACTGCTGAGCACTGCTGAGCACTGCTG 120
DB AGCAATGAATTCATCAATTCATGAGCACTGCTGAGCACTGCTGAGCACTGCTG 186
QY 121 TCCAAAGCACTCTTCCAAAGCACTGCTGAGCACTGCTGAGCACTGCTGAGCACT 180
DB TCCAAAGCACTCTTCCAAAGCACTGCTGAGCACTGCTGAGCACTGCTGAGCACT 246
QY 181 CACTGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB CACTGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 306
QY 241 AAGGCGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACT 300

Db 307 AAGGCCAGCATGACACATGAGCGGCGCCCTGCGCTGGAACCTGCACTGTCCTT 366
Qy 301 CAGCAACGCTACCATGCCCCACAGATCTGATGCTCTTCACTGGGCTGGGAAACATAT 360
Db 367 CAGCAACGCTACCATGCCCCACAGATCTGATGCTCTTCACTGGGCTGGGAAACATAT 426
Qy 361 TACTGAGGAGCCAGACACCGGAGGAGCCCTGCTGATGTCAGAGTGGGCTTAAG 420
Db 427 TACTGAGGAGCCAGACACCGGAGGAGCCCTGCTGATGTCAGAGTGGGCTTAAG 486
Qy 421 CGCTTGTCCAAAGATGATGTCATGTCGAGAGATGAAAAAAGCCCTCCCTCCT 480
Db 487 CGCTTGTCCAAAGATGATGTCATGTCGAGAGATGAAAAAAGCCCTCCCTCCT 546
Qy 481 CCAGAAAGATTAATTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
Db 547 CCAGAAAGATTAATTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 606
Qy 541 GGGGGAATTCACCAACCATGAGAACGAGCCCTGTTTGCGGCATCTACAGAGGCAC 600
Db 607 GGGGGAATTCACCAACCATGAGAACGAGCCCTGTTTGCGGCATCTACAGAGGCAC 666
Qy 601 GGGGGGGCTCTGTCACTTACCTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 667 GGGGGGGCTCTGTCACTTACCTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 726
Qy 661 AGGCGACACACTGCTTCACTGATTTACCCAAAGAGAGGAGGAGGAGGAGGAGGAGG 720
Db 727 AGGCGACACACTGCTTCACTGATTTACCCAAAGAGAGGAGGAGGAGGAGGAGGAGG 786
Qy 721 CGCTCAAGGCTTAACTCCAAACGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 787 CGCTCAAGGCTTAACTCCAAACGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 846
Qy 781 CTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 847 CTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
Qy 841 ATCCGTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db 907 ATCCGTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 966
Qy 901 CCTCTGATGTAATTCATCCCAAGTTGGGACAAAGCTGTGAGATCACTGGCTTTGAAA 960
Db 967 CCTCTGATGTAATTCATCCCAAGTTGGGACAAAGCTGTGAGATCACTGGCTTTGAAA 1026
Qy 961 GAGAAATCTACCGGATCTCTATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
Db 1027 GAGAAATCTACCGGATCTCTATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1086
Qy 1021 TCCCAACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 1087 TCCCAACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1146
Qy 1081 TGTGCTGCTACCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1147 TGTGCTGCTACCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1206
Qy 1141 GTCTGTTCCTTCCAAAGGCGCATGACTTTGATCTGGAATTTGAGTGGGGCGGTGAGAT 1200
Db 1207 GTCTGTTCCTTCCAAAGGCGCATGACTTTGATCTGGAATTTGAGTGGGGCGGTGAGAT 1266
Qy 1201 GGCCTGAAGAGCAAGCCAGGCGTCTACAGAGAGTCTCAACTTTACCTGATCCGCG 1260
Db 1267 GGCCTGAAGAGCAAGCCAGGCGTCTACAGAGAGTCTCAACTTTACCTGATCCGCG 1326
Qy 1261 AGTCAACCAAGAGAGGAGATGGCGTGGCGCTCTGA 1296
Db 1327 AGTCAACCAAGAGAGGAGATGGCGTGGCGCTCTGA 1362

US-09-480-884A-123
; Sequence 123, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455c6
; CURRENT APPLICATION NUMBER: US/09/480, 884A
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-123

Query Match 100.0%; Score 1296; DB 4; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGGCCCTGCTGCGCGGCGCTGCTCTGCGTCCTGTCGAGGAGGAGGAGGAGGAGG 60
Db 67 ATGAGGCCCTGCTGCGCGGCGGCGCTGCTCTGCGTCCTGTCGAGGAGGAGGAGGAGG 126
Qy 61 AGCAATGAACCTTCAATCAAGTTCCATGGAACCTGTCATCTTAATGAGAGAACATGTGT 120
Db 127 AGCAATGAACCTTCAATCAAGTTCCATGGAACCTGTCATCTTAATGAGAGAACATGTGT 186
Qy 121 TCCAAAGATCTTCTCCAAATTCATCTGTCGAACTGCCAAAGAAATTCGAGGGGAG 180
Db 187 TCCAAAGATCTTCTCCAAATTCATCTGTCGAACTGCCAAAGAAATTCGAGGGGAG 246
Qy 181 CACTGTGAATATGATTAATCAAAAACCTGCTATGAGGGGAAATGTCACCTTTTACGAGGA 240
Db 247 CACTGTGAATATGATTAATCAAAAACCTGCTATGAGGGGAAATGTCACCTTTTACGAGGA 306
Qy 241 AAGGCCACACCTGACACCATGAGGCGGCGGCGCTGCGCTGGAACTCTGCGACATGTCCT 300
Db 307 AAGGCCACACCTGACACCATGAGGCGGCGGCGCTGCGCTGGAACTCTGCGACATGTCCT 366
Qy 367 CAGCAACGTAACCATGCCCACAGATCTGATGCTCTTCAAGCTGGGCTGGGGAAACATAT 426
Db 427 TACTGAGGAGCCAGACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 486
Qy 421 CGGCTTGTCCAAAGATGATGTCATGTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Db 487 CGGCTTGTCCAAAGATGATGTCATGTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 546
Qy 481 CCAGAAAGATTAATTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
Db 547 CCAGAAAGATTAATTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 606
Qy 541 GGGGGAATTCACCAACCATGAGAACGAGCCCTGTTTGCGGCATCTACAGAGGCAC 600
Db 607 GGGGGAATTCACCAACCATGAGAACGAGCCCTGTTTGCGGCATCTACAGAGGCAC 666
Qy 601 GGGGGGGCTCTGTGACCTTACCTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 667 GGGGGGGCTCTGTGACCTTACCTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 726
Qy 661 AGGCGACACACTGCTTCACTGATTTACCCAAAGAGAGGAGGAGGAGGAGGAGGAGG 720
Db 727 AGGCGACACACTGCTTCACTGATTTACCCAAAGAGAGGAGGAGGAGGAGGAGGAGG 786

OY	721	CGCTCAGAGCTTAACTACCAACCGGAGAGATGAACTTTAGAGTGTGAAAAACCTATC	780
Db	787	CGCTCAGAGCTTAACTACCAACCGGAGAGATGAACTTTAGAGTGTGAAAAACCTATC	846
OY	781	CTACACAGAGACTACAGGCTGACAGGCTGTGTCCACCAAGACATTCGCTTGTAAG	840
Db	847	CTACACAGAGACTACAGGCTGACAGGCTGTGTCCACCAAGACATTCGCTTGTAAG	906
OY	841	ATCCGTTCCAGAGAGGAGGTGTGTGCGAGCCATCCCGAGTATACAGACATCTGCTG	900
Db	907	ATCCGTTCCAGAGAGGAGGTGTGTGCGAGCCATCCCGAGTATACAGACATCTGCTG	966
OY	901	CCCTCGATGATATACGATCCCAAGTTTGGACCAAGCTGTGATCATCTGGCTTTGAAAA	960
Db	967	CCCTCGATGATATACGATCCCAAGTTTGGACCAAGCTGTGATCATCTGGCTTTGAAAA	1026
OY	961	GAGAAATTCATCCGACTATCTATCCCGAGAGAGCTGAAAAATGACTGTGTGAACTGATT	1020
Db	1027	GAGAAATTCATCCGACTATCTATCCCGAGAGAGCTGAAAAATGACTGTGTGAACTGATT	1086
OY	1021	TCCCAACCGGAGTGTACAGAGCCCACTACTACGGCTGTGAAGTCAACCAAAATGCTG	1080
Db	1087	TCCCAACCGGAGTGTACAGAGCCCACTACTACGGCTGTGAAGTCAACCAAAATGCTG	1146
OY	1081	TGTGTCTGTGACCCACAGTGTGAAAAAGATTCTTGCCAGGAGACTCAGGGGAGCCCTC	1140
Db	1147	TGTGTCTGTGACCCACAGTGTGAAAAAGATTCTTGCCAGGAGACTCAGGGGAGCCCTC	1206
OY	1141	GTCGTTCCTCTCCAAAGCCCGCATGACTTTGACTGTGAATTGTGAGCTGTGGGCCGTGATGT	1200
Db	1207	GTCGTTCCTCTCCAAAGCCCGCATGACTTTGACTGTGAATTGTGAGCTGTGGGCCGTGATGT	1266
OY	1201	GCCCTGAGAGCAAGCCAGGGGCTCAACAGAGAGTCCACACTTCTTAACCTGATCCG	1260
Db	1267	GCCCTGAGAGCAAGCCAGGGGCTCAACAGAGAGTCCACACTTCTTAACCTGATCCG	1326
OY	1261	AGTCACACCAAGAGAGATGGCTCGGACCTCTGA	1296
Db	1327	AGTCACACCAAGAGAGATGGCTCGGACCTCTGA	1362

RESULT 3
US-09-542-615A-123
; Sequence 123. Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hoshen, Nancy A.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542.615A
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-123

Query Match	Similarity	Score	DB 4	Length	2254
Best Local	Similarity	100.0%	Pred. No. 0		
Match	Conservative	100.0%	Mismatches	0	
			Indels	0	
			Gaps	0	
Qy	1	ATGAGAGCCCTGCTGCGCGCCCTGCTTCTTCGCTCTGATCGAGCGACTCCAAAGC	60		
Db	67	ATGAGAGCCCTGCTGCGCGCCCTGCTTCTTCGCGCTCGATCGAGCGACTCCAAAGC	126		

QY	61	AGCATGAACCTTCATCAAGTTCCATCCGACCTGTGACTCTCTTAAATGGAGGAAACATGTG	120
Db	127	AGCAATGAACCTTCATCAAGTTCCATCCGACCTGTGACTCTCTTAAATGGAGGAAACATGTG	186
QY	121	TCCAACAAGTACTCTCCCAATTCACCTGTGTCAACTGCCCCAAGAAATTCGAGGGCAG	180
Db	187	TCCAACAAGTACTCTCCCAATTCACCTGTGTCAACTGCCCCAAGAAATTCGAGGGCAG	246
QY	181	CACGTGAAATGATTAAGTCAAAAACCTGCTATGAGGGAAATGGTCACTTTTACCGAGA	240
Db	247	CACGTGAAATGATTAAGTCAAAAACCTGCTATGAGGGAAATGGTCACTTTTACCGAGA	306
QY	241	AAGGCCACACTGACACCATGGGCGGCGCCCTGCGCTCTGGAATCTGCCACTGTCCCT	300
Db	307	AAGGCCACACTGACACCATGGGCGGCGCCCTGCGCTCTGGAATCTGCCACTGTCCCT	366
QY	301	CAGCAAAAGTACATGCCCAAGATCTGATGCTCTTCAGCTGGGCTGGGGAACATAAT	360
Db	367	CAGCAAAAGTACATGCCCAAGATCTGATGCTCTTCAGCTGGGCTGGGGAACATAAT	426
QY	361	TACTGCAGAAACCCAGACAAACCGGAGGGGACCCCTGTGTCTATGTGCAGGTGGGCTTAAG	420
Db	427	TACTGCAGAAACCCAGACAAACCGGAGGGGACCCCTGTGTCTATGTGCAGGTGGGCTTAAG	486
QY	421	CGCGTTGTCCAAAGAGTGAATGGTGCATACTGCGCAGATGGAATAAAACCCCTCTCCT	480
Db	487	CGCGTTGTCCAAAGAGTGAATGGTGCATACTGCGCAGATGGAATAAAACCCCTCTCCT	546
QY	481	CCAGAAATTTAAATTTCACTGTGGCCAAAGAATCTGAGGCCCCGCTTTAAGTATTT	540
Db	547	CCAGAAATTTAAATTTCACTGTGGCCAAAGAATCTGAGGCCCCGCTTTAAGTATTT	606
QY	541	GGGGGAGAATTCACCCACCATGAGAAACCAAGCCCTGTGTTGGGGCATTTACAGAGGAC	600
Db	607	GGGGGAGAATTCACCCACCATGAGAAACCAAGCCCTGTGTTGGGGCATTTACAGAGGAC	666
QY	601	CGGGGGGCTCTGTACACCTAGCTGTGTGTGAGAGCAGCTCATCAGCCCTGTGGGTGATC	660
Db	667	CGGGGGGCTCTGTACACCTAGCTGTGTGTGAGAGCAGCTCATCAGCCCTGTGGGTGATC	726
QY	661	AGCGGCACACCTGCTTATGATTTACCCAAAGAAAGAGACTCATATGTCTTACTGGGT	720
Db	727	AGCGGCACACCTGCTTATGATTTACCCAAAGAAAGAGACTCATATGTCTTACTGGGT	786
QY	721	CGCTCAAGGCTTAATCCCAACGCGAAGGGAGATGAATTTAGAGTGGAAACCTAATC	780
Db	787	CGCTCAAGGCTTAATCCCAACGCGAAGGGAGATGAATTTAGAGTGGAAACCTAATC	846
QY	781	CTACACAGAGACTACAGGCTGACACGCTTCTCACCAACGACATTTGCTGTGAAG	840
Db	847	CTACACAGAGACTACAGGCTGACACGCTTCTCACCAACGACATTTGCTGTGAAG	906
QY	841	ATCGCTTCCAAAGGAGGGAGGTGTGCGAGCAATCCCGAGCTTATACACATCTGCCTG	900
Db	907	ATCGCTTCCAAAGGAGGGAGGTGTGCGAGCAATCCCGAGCTTATACACATCTGCCTG	966
QY	901	CCCTGATGTATTAACATCCCAAGTTTGGCAACAAGCTGTGATCTACTGCTTTGGAAAA	960
Db	967	CCCTGATGTATTAACATCCCAAGTTTGGCAACAAGCTGTGATCTACTGCTTTGGAAAA	1026
QY	961	GAGAAATTTACCGACTATCTTATCCGAGCAGCTGAAAAATGACTGTGTGTGAAGTAAAT	1020
Db	1027	GAGAAATTTACCGACTATCTTATCCGAGCAGCTGAAAAATGACTGTGTGTGAAGTAAAT	1086
QY	1021	TCCCAACGGGAGTGTACAGCAGCCCACTACTACGGCTGTGAAGTCAACACAAAATGCTG	1080
Db	1087	TCCCAACGGGAGTGTACAGCAGCCCACTACTACGGCTGTGAAGTCAACACAAAATGCTG	1146
QY	1081	TGTGTGCTGACCCACAGTGGAAAAACAGATTTCTGTCCAGGGAGACTCAGGGGAGCCCTC	1140
Db	1147	TGTGTGCTGACCCACAGTGGAAAAACAGATTTCTGTCCAGGGAGACTCAGGGGAGCCCTC	1206

[illegible]

STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023.655
FILING DATE: HEREMITH
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1217:
SEQUENCE CHARACTERISTICS:
LENGTH: 2294 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g220138
US-09-023-655-1217

Query Match 100.0%; Score 1296; DB 4; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGGCGCCCTGCTTCTGCGTCTGCTGCGAGCACTCCAAAGC 60
DB 67 ATGAGAGCCCTGCTGGCGCCCTGCTTCTGCGTCTGCTGCGAGCACTCCAAAGC 126
QY 61 AGCAATGAATTTATCAAGTTCATGCACTGCACTGCTGCTTAATGAGAGCAATGTGTG 120
DB 127 AGCAATGAATTTATCAAGTTCATGCACTGCACTGCTGCTTAATGAGAGCAATGTGTG 186
QY 121 TCCAAAGATCTTCCCAATCTGCTGCACTGCTGCACTGCTGCACTGCTGCACTGCTGCA 180
DB 187 TCCAAAGATCTTCCCAATCTGCTGCACTGCTGCACTGCTGCACTGCTGCACTGCTGCA 246
QY 181 CACTGGAATATAGTCAATGCTCAAAACCTGCTATGAGGGAATGCTCACTTTACGAGGA 240
DB 247 CACTGGAATATAGTCAATGCTCAAAACCTGCTATGAGGGAATGCTCACTTTACGAGGA 306
QY 241 AAGGCGAGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCTT 300
DB 307 AAGGCGAGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCTT 366
QY 301 CAGCAAGATGATGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCTT 360
DB 367 CAGCAAGATGATGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCTT 426
QY 361 TACTGAGGAAACCCAGAGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCTT 420
DB 427 TACTGAGGAAACCCAGAGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCTT 486
QY 421 CCGCTGTGCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 480
DB 487 CCGCTGTGCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 546
QY 481 CCGAGAGATTAATTTTCAAGTGTGCGCAAAAGACTGAGGCGCCGCTTAAATTTAT 540

DB 547 CCGAGAGATTAATTTTCAAGTGTGCGCAAAAGACTGAGGCGCCGCTTAAATTTAT 606
QY 541 GGGGAGAGATTTACCAACCATGAGAACCAAGCCCTGCTTGGCCATCTACAGAGGCA 600
DB 607 GGGGAGAGATTTACCAACCATGAGAACCAAGCCCTGCTTGGCCATCTACAGAGGCA 666
QY 601 GGGGAGAGATTTACCAACCATGAGAACCAAGCCCTGCTTGGCCATCTACAGAGGCA 660
DB 667 GGGGAGAGATTTACCAACCATGAGAACCAAGCCCTGCTTGGCCATCTACAGAGGCA 726
QY 661 AGGCGCACACACTGCTTCAATTTGATTAACCAAGAGAGGAGTCACTGCTGCTGCTG 720
DB 727 AGGCGCACACACTGCTTCAATTTGATTAACCAAGAGAGGAGTCACTGCTGCTGCTG 786
QY 721 CGCTCAAGGCTTAATCTCAACCAAGAGAGGAGTCACTGCTGCTGCTGCTGCTG 780
DB 787 CGCTCAAGGCTTAATCTCAACCAAGAGAGGAGTCACTGCTGCTGCTGCTGCTGCTG 846
QY 781 CTACAGAGATTTACCAAGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 847 CTACAGAGATTTACCAAGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
QY 841 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCCGGACTATACAGACATCTGCTG 900
DB 907 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCCGGACTATACAGACATCTGCTG 966
QY 901 CCCTGATGTATACGATCCCAAGTTTGGCAAGGCTGTGAGATCACTGCTTGGAAAA 960
DB 967 CCCTGATGTATACGATCCCAAGTTTGGCAAGGCTGTGAGATCACTGCTTGGAAAA 1026
QY 961 GAGAAATTTACCAAGATTTCTATCTGCTGAGAGAGTGAATGATCTGTTGGAAGCTGAT 1020
DB 1027 GAGAAATTTACCAAGATTTCTATCTGCTGAGAGAGTGAATGATCTGTTGGAAGCTGAT 1086
QY 1021 TCCCAAGGAGATGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 1087 TCCCAAGGAGATGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1146
QY 1081 TGTGCTGTGACCCCAAGTGAAGAAAGATTCCTGCGAGGAGATCTAGAGGAGGAGG 1140
DB 1147 TGTGCTGTGACCCCAAGTGAAGAAAGATTCCTGCGAGGAGATCTAGAGGAGGAGG 1206
QY 1141 GTCTGTTCCCTCCAAAGGCGGATGATTTGATGCTGGAATTTGTGAGCTGGGCGGTGAT 1200
DB 1207 GTCTGTTCCCTCCAAAGGCGGATGATTTGATGCTGGAATTTGTGAGCTGGGCGGTGAT 1266
QY 1201 GCCCTGAAGAGCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1267 GCCCTGAAGAGCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1326
QY 1261 AGTCACCAAGAGAGAGAGATGCTGAGGCTTGA 1296
DB 1327 AGTCACCAAGAGAGAGATGCTGAGGCTTGA 1362

RESULT 6
US-09-221-107-123
Sequence 123, Application US/09221107
Patent No. 6660838
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C2
CURRENT APPLICATION NUMBER: US/09/221.107
CURRENT FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 161
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 123
LENGTH: 2294
TYPE: DNA
ORGANISM: Human
US-09-221-107-123

QY 301 CAGCAACGACATGCCCCAGATCTGATGCTTTCAGCTGGCCCTGGGAAACATTAAT 360
 DB 381 CAGCAACGACATGCCCCAGATCTGATGCTTTCAGCTGGCCCTGGGAAACATTAAT 440
 QY 361 TACTGAGGAAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
 DB 441 TACTGAGGAAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 500
 QY 421 CCGCTGTCCAGAGTGCATGTGTGATGATGCTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
 DB 501 CCGCTGTCCAGAGTGCATGTGTGATGATGCTGCGCAGATGGAAGAAAGCCCTCTCTCT 560
 QY 481 CCGAAGAAATTAATTTTGTGTGTGCGCAAAACATCTGAGGCGCCCTTAAGATTAAT 540
 DB 561 CCGAAGAAATTAATTTTGTGTGTGCGCAAAACATCTGAGGCGCCCTTAAGATTAAT 620
 QY 541 GGGGAGAAATTCACCAACATCGAGAACACAGCCCTGTTTGGCCATCTACAGAGGAC 600
 DB 621 GGGGAGAAATTCACCAACATCGAGAACACAGCCCTGTTTGGCCATCTACAGAGGAC 680
 QY 601 CCGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGAGCCCTATCAGCCCTTGCGGTGATC 660
 DB 681 CCGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGAGCCCTATCAGCCCTTGCGGTGATC 740
 QY 661 AGGCGCACACACTGCTTCAATTAATTAACCAAGAGAGGACTACATCTGCTACCTGGT 720
 DB 741 AGGCGCACACACTGCTTCAATTAATTAACCAAGAGAGGACTACATCTGCTACCTGGT 800
 QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGATGAGTGAAGTGAAGAAACCTAATC 780
 DB 801 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGATGAGTGAAGTGAAGAAACCTAATC 860
 QY 781 CTACACAGAGACTACAGCGCTGACACGCTTGTCTACACACAGCAATGCTGCTGAG 840
 DB 861 CTACACAGAGACTACAGCGCTGACACGCTTGTCTACACACAGCAATGCTGCTGAG 920
 QY 841 ATCCGCTTCCAGAGAGGAGGAGTGTGGGAGCCATCCCGGACTATACAGACATCTGCTG 900
 DB 921 ATCCGCTTCCAGAGAGGAGGAGTGTGGGAGCCATCCCGGACTATACAGACATCTGCTG 980
 QY 901 CCTCGATGATTAACGATCCCGAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAA 960
 DB 981 CCTCGATGATTAACGATCCCGAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAA 1040
 QY 961 GAGAAATTCACGCACTATCTATCCGAGACAGCTGAAATGACTGTTGTGAGCTGATT 1020
 DB 1041 GAGAAATTCACGCACTATCTATCCGAGACAGCTGAAATGACTGTTGTGAGCTGATT 1100
 QY 1021 TCCCAACCGGAGAGTGTGACAGCCCGCACTACTACGCTCTGAAATCACACAAATGCTG 1080
 DB 1101 TCCCAACCGGAGAGTGTGACAGCCCGCACTACTACGCTCTGAAATCACACAAATGCTG 1160
 QY 1081 TGTGTGTGTGACCCAGAGTGAAGAAACAGATTCTGTGACGAGGAGACTCAGGGGAGACCCCT 1140
 DB 1161 TGTGTGTGTGACCCAGAGTGAAGAAACAGATTCTGTGACGAGGAGACTCAGGGGAGACCCCT 1220
 QY 1141 GTCTGTTCCTCCAGAGCGCATGACTTTGACTGGAATGTGAGCTGGGCGCTGTGATG 1200
 DB 1221 GTCTGTTCCTCCAGAGCGCATGACTTTGACTGGAATGTGAGCTGGGCGCTGTGATG 1280
 QY 1201 GCCCTGAGAGACAGCGAGCGCTGTACAGAGAGTCTCACACTTCTTAACCTGGATCCG 1260
 DB 1281 GCCCTGAGAGACAGCGAGCGCTGTACAGAGAGTCTCACACTTCTTAACCTGGATCCG 1340
 QY 1261 AGTCAACCAAGAGAGAGATGGGCTGGCCCTCTGA 1296
 DB 1341 AGTCAACCAAGAGAGAGATGGGCTGGCCCTCTGA 1376

RESULT 8
 US-09-480-884A-122
 ; Sequence 122, Application US/09480884A

; Patent No. 6482597
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LONG CANCER
 ; FILE REFERENCE: 210121.455C6
 ; CURRENT APPLICATION NUMBER: US/09/480,884A
 ; CURRENT FILING DATE: 2001-08-27
 ; NUMBER OF SEQ ID NOS: 330
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 122
 ; LENGTH: 1475
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-480-884A-122

Query Match 99.8%; Score 1292.8; DB 4; Length 1475;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGGCGGGCCCTGCTTCTGTGCTCTGTGTGAGGCGACTCCAAAGC 60
 DB 81 ATGAGAGCCCTGCTGGCGGGCCCTGCTTCTGTGCTCTGTGTGAGGCGACTCCAAAGC 140
 QY 61 AGCAATGAATCTTCATCAAGTTCATGCACTGCACTGTGCTTAATGAGAGAACTATGTG 120
 DB 141 AGCAATGAATCTTCATCAAGTTCATGCACTGCACTGTGCTTAATGAGAGAACTATGTG 200
 QY 121 TCCAAACAAGTACTTCTTCCAAACATTCACGTGTGCACTGCCCAAGAAATTCGAGGCG 180
 DB 201 TCCAAACAAGTACTTCTTCCAAACATTCACGTGTGCACTGCCCAAGAAATTCGAGGCG 260
 QY 181 CACTGTGAATAGTAAAGTCAAAACCTGCTGTAAGAGGGAGATGGTCACTTTTACCGAGGA 240
 DB 261 CACTGTGAATAGTAAAGTCAAAACCTGCTGTAAGAGGGAGATGGTCACTTTTACCGAGGA 320
 QY 241 AAGGCCAGACTGACACACATGAGCGCCGAGCCCTGCTGCTGCAACTCTGCACTGTCTT 300
 DB 321 AAGGCCAGACTGACACACATGAGCGCGCCGAGCCCTGCTGCTGCAACTCTGCACTGTCTT 380
 QY 301 CAGCAACGTAACCATCCCAACAGATGTGATGCTTCACTGAGGCTTGGGAAACATTAAT 360
 DB 381 CAGCAACGTAACCATCCCAACAGATGTGATGCTTCACTGAGGCTTGGGAAACATTAAT 440
 QY 361 TACTGAGGAAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
 DB 441 TACTGAGGAAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 500
 QY 421 CCGCTGTCCAGAGTGCATGTGTGATGATGCTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
 DB 501 CCGCTGTCCAGAGTGCATGTGTGATGATGCTGCGCAGATGGAAGAAAGCCCTCTCTCT 560
 QY 481 CCGAAGAAATTAATTTTGTGTGTGCGCAAAACATCTGAGGCGCCCTTAAGATTAAT 540
 DB 561 CCGAAGAAATTAATTTTGTGTGTGCGCAAAACATCTGAGGCGCCCTTAAGATTAAT 620
 QY 541 GGGGAGAAATTCACCAACATCGAGAACACAGCCCTGTTTGGCCATCTACAGAGGAC 600
 DB 621 GGGGAGAAATTCACCAACATCGAGAACACAGCCCTGTTTGGCCATCTACAGAGGAC 680
 QY 601 CCGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGAGCCCTATCAGCCCTTGCGGTGATC 660
 DB 681 CCGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGAGCCCTATCAGCCCTTGCGGTGATC 740
 QY 661 AGGCGCACACACTGCTTCAATTAATTAACCAAGAGAGGACTACATCTGCTACCTGGT 720
 DB 741 AGGCGCACACACTGCTTCAATTAATTAACCAAGAGAGGACTACATCTGCTACCTGGT 800
 QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGATGAGTGAAGTGAAGAAACCTAATC 780

Db	801	CGTCTAAGGCTTAACTTCCAAACAGCAAGGGGAATGATGTAAGGTGGAAAACTTCATC	860
Qy	781	CTACACAAAGACTTACAGCGCTGACAGCGTTGCTCACACACAGACATTCGCTTGCGAAG	840
Db	861	CTACACAAAGACTTACAGCGCTGACAGCGTTGCTCACACACAGACATTCGCTTGCGAAG	920
Qy	841	ATCCGTTTCCAAAGAGGGCAGGTGTGGCGACCATCCCGGACTATACAGACCATCTGCGTG	900
Db	921	ATCCGTTTCCAAAGAGGGCAGGTGTGGCGACCATCCCGGACTATACAGACCATCTGCGTG	980
Qy	901	CCCTGATGTATTAACGATCCCGAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAA	960
Db	981	CCCTGATGTATTAACGATCCCGAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAA	1040
Qy	961	GAGAAATTTAAACGACTATCTTATTCGGAGCAGCTGAAAATGACTGTGTGAAGCTGAT	1020
Db	1041	GAGAAATTTAAACGACTATCTTATTCGGAGCAGCTGAAAATGACTGTGTGAAGCTGAT	1100
Qy	1021	TCCCAACCGGAGGTGTAGAGGCCCACTTACTTACCGGAGCAGCTGAAGTCAACCAAAATGCTG	1080
Db	1101	TCCCAACCGGAGGTGTAGAGGCCCACTTACTTACCGGAGCAGCTGAAGTCAACCAAAATGCTG	1160
Qy	1081	TGTGCTGTGACCCACAGTGGAAAACAGATTCTCTGCAAGGAGACTCAAGGSGGACCCCTC	1140
Db	1161	TGTGCTGTGACCCACAGTGGAAAACAGATTCTCTGCAAGGAGACTCAAGGSGGACCCCTC	1220
Qy	1141	GTCTGTTCCTCTCAAGGCCGCGATGACTTTGACCTGGAATTGTGAGCTGGGGCGCTGAGATGT	1200
Db	1221	GTCTGTTCCTCTCAAGGCCGCGATGACTTTGACCTGGAATTGTGAGCTGGGGCGCTGAGATGT	1280
Qy	1201	GCCCTGAAGAGACAAACCGAGCGCTCTACACAGAGCTGCACATTCCTTACCCCGAATCCG	1260
Db	1281	GCCCTGAAGAGACAAACCGAGCGCTCTACACAGAGCTGCACATTCCTTACCCCGAATCCG	1340
Qy	1261	AGTCACACCAAGAAAGAGATGCGCTGGCCCTCTGA	1296
Db	1341	AGTCACACCAAGAAAGAGATGCGCTGGCCCTCTGA	1376

```

RESULT 9
US-09-542-615A-122
Sequence 122, Application US/09542615A
Patent No. 6518256
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Lijun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaltanya S.
APPLICANT: Hosken, Nancy A.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.45568
CURRENT APPLICATION NUMBER: US/09/542,615A
CURRENT FILING DATE: 2000-04-14
NUMBER OF SEQ. ID NOS: 350
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 122
LENGTH: 1475
TYPE: DNA
ORGANISM: Homo sapien
US-09-542-615A-122

```

Query Match	99.8%	Score 1292.8	DB 4	Length 1475
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1294	Conservative	0	Mismatches 2	Indels 0
			Gaps	0

QY 1 ATGAGAGCCCTGCTGGCGGCCCTGCTTCCTGCGCTGCTGCTGAGGACTCCAAAGC 60
Db 81 ATGAGAGCCCTGCTGGCGGCCCTGCTTCCTGCGCTGCTGCTGAGGACTCCAAAGC 140
QY 61 AGCAATGAACTTATTAAGTTCATCCATCGAAGCTGTGACTCTTAAATGGAGAACTGTGG 120

Db	141	AGCAATGAACTTATCAAGTTCCATGAACTGAACTGCTTAATATGAGAAACATGTGTG	200
QY	121	TCGAACAAGTACTTCTCCAACTTCACTGCTGTCMACTGCCCCAAAGAAATTCGAGGGCAG	180
Db	201	TCCAAACAAGTACTTCTCCAACTTCACTGCTGTCMACTGCCCCAAAGAAATTCGAGGGCAG	260
QY	181	CAGTGTGAAATTAATATAGTCAAAAACCTGCTATAGAGGGGAATGTGACTTTTACGAGGA	240
Db	261	CAGTGTGAAATTAATATAGTCAAAAACCTGCTATAGAGGGGAATGTGACTTTTACGAGGA	320
QY	241	AAGCCAGCACTGACACATGAGGCGGGCCCTGCTCCCTGGAACCTGTGCACATGTCTT	300
Db	321	AAGCCAGCACTGACACATGAGGCGGGCCCTGCTCCCTGGAACCTGTGTCCACTGTCTT	380
QY	301	CAGCAACGTAACATGCCCCACAGATCTGATGCTCTTCACTGAGGCTGTGGGAAACATAAT	360
Db	381	CAGCAACGTAACATGCCCCACAGATCTGATGCTCTTCACTGAGGCTGTGGGAAACATAAT	440
QY	361	TACTGAGGAACCCAGACAACCGAGGGGCAACCTGTGTGTATAGTGAAGTGGGCTTAAG	420
Db	441	TACTGAGGAACCCAGACAACCGAGGGGCAACCTGTGTGTATAGTGAAGTGGGCTTAAG	500
QY	421	CCGCTTGTCCAAAGTGCATGTGTGCATGACTCGCAGATGAAAAAGCCCTCTCTCT	480
Db	501	CCGCTTGTCCAAAGTGCATGTGTGCATGACTCGCAGATGAAAAAGCCCTCTCTCTCT	560
QY	481	CCAGAAAGATTTAAATTTCAATGTGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTAAT	540
Db	561	CCAGAAAGATTTAAATTTCAATGTGTGGCCAAAGAAGCTCTGAGGCCCGCTTTAAGATTAAT	620
QY	541	GGGGGAGATTCACACATCGAGAACCGCCCTGTTTGGCGGCATCTACAGGAGGCAC	600
Db	621	GGGGGAGATTCACACATCGAGAACCGCCCTGTTTGGCGGCATCTACAGGAGGCAC	680
QY	601	CGGGGGGGGCTCTGTCACTTACGTGTGTGAGGCAAGCTTCATCAAGCCCTTGTGGGTGATC	660
Db	681	CGGGGGGGGCTCTGTCACTTACGTGTGTGAGGCAAGCTTCATCAAGCCCTTGTGGGTGATC	740
QY	661	AGCGGCACACACTGCTTACATTGATTAACCAAGGAAGGAGACTAATGTCTACCTGGGT	720
Db	741	AGCGGCACACACTGCTTACATTGATTAACCAAGGAAGGAGACTAATGTCTACCTGGGT	800
QY	721	CGCTCAAGGCTTAACCTCCAAACCGCAAGGAGGATGAAGTTTGAAGTGAAGAACTAATC	780
Db	801	CGCTCAAGGCTTAACCTCCAAACCGCAAGGAGGATGAAGTTTGAAGTGAAGAACTAATC	860
QY	781	CTACACAAGGCTACAGGCTGACACGCTTGTCTACCAACAAGCATTTGCTTGTCTGAAG	840
Db	861	CTACACAAGGCTACAGGCTGACACGCTTGTCTACCAACAAGCATTTGCTTGTCTGAAG	920
QY	841	ATCCGTTTCCAAAGAGGAGGAGGTGTGTGCGAGGCATTCGCCGACTATACAGACCATGTGCTG	900
Db	921	ATCCGTTTCCAAAGAGGAGGAGGTGTGTGCGAGGCATTCGCCGACTATACAGACCATGTGCTG	980
QY	901	CCCTGTGATGTAAAGATCCCAAGTTTGGCAACAAGCTGTGAGATCACTGCTTTGGAAA	960
Db	981	CCCTGTGATGTAAAGATCCCAAGTTTGGCAACAAGCTGTGAGATCACTGCTTTGGAAA	1040
QY	961	GAGAAATTCTACGACTATCTATTCGAGAGAGCTGAAAAAGACTGTTTGTGAACTGATTT	1020
Db	1041	GAGAAATTCTACGACTATCTATTCGAGAGAGCTGAAAAAGACTGTTTGTGAACTGATTT	1100
QY	1021	TCCCAACGGGAGGTGCAGACGCCCACTACTTAACGCTCTGTGAAGTCAACCAACAAATGCTG	1080
Db	1101	TCCCAACGGGAGGTGCAGACGCCCACTACTTAACGCTCTGTGAAGTCAACCAACAAATGCTG	1160
QY	1081	TGTGTCTGTGACCCACAGTGGAAAAAGATTCTGTCCAGGAGAGACTCAGGGGGGACCCCTC	1140
Db	1161	TGTGTCTGTGACCCACAGTGGAAAAAGATTCTGTCCAGGAGAGACTCAGGGGGGACCCCTC	1220
QY	1141	GTCGTGTTCCCTTCAAGGCGCGATACCTTTGACTGTGAATTGTGAGCTGTGGGCGTGAAGT	1200

Db 1221 GTCGTGCTCCCTCCAGGCGCCGATGATCTTGATGGAATGAGCTGGGCGCGTGATGT 1280
Qy 1201 GCCCTGAAGGACAGAGCGGCTCTACAGAGTCTCAACATTCCTTACCTTGATGATCCG 1260
Db 1281 GCCCTGAAGGACAGAGCGGCTCTACAGAGTCTCAACATTCCTTACCTTGATGATCCG 1340
Qy 1261 AGTCACACCAAGAGAGATGAGCTGGGCGCGTGATGT 1296
Db 1341 AGTCACACCAAGAGAGATGAGCTGGGCGCGTGATGT 1376
RESULT 10
US-09-606-421B-122
; Sequence 122, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606.421B
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-122
Query Match 99.8%; Score 1292.8; DB 4; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGAGCCCTGCTGCGCGCCCTGCTCTCTGCTGCTGCTGCGAGCACTCCAAAGC 60
Db 81 ATGAGAGCCCTGCTGCGCGCCCTGCTCTCTGCTGCTGCTGCGAGCACTCCAAAGC 140
Qy 61 AGCAATGAATCTCATCAAGTTCATCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 141 AGCAATGAATCTCATCAAGTTCATCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 200
Qy 121 TCCAAAGAGTACTTCTCCAAATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 201 TCCAAAGAGTACTTCTCCAAATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 260
Qy 181 CACTGTGAATATGATTAAGTCAAAACCTGCTATGAGGAGATGCTGCTGCTGCTGCTGCT 240
Db 261 CACTGTGAATATGATTAAGTCAAAACCTGCTATGAGGAGATGCTGCTGCTGCTGCTGCT 320
Qy 241 AAGGCGAGCACTGACACATGAGGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 321 AAGGCGAGCACTGACACATGAGGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
Qy 301 CAGCAAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 381 CAGCAAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440
Qy 361 TACTGAGGAAACCAAGACCGGAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 441 TACTGAGGAAACCAAGACCGGAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
Qy 421 CCGCTTGTCCAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 501 CCGCTTGTCCAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560

Qy 481 CCAGAAAGATTAATTAATTTGAGTGGCCAAAGAGCTGTAGGGCCGCTTTAAGATTAT 540
Db 561 CCAGAAAGATTAATTAATTTGAGTGGCCAAAGAGCTGTAGGGCCGCTTTAAGATTAT 620
Qy 541 GGGGGAAGATTTCAACCAATCGAAGAACCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 621 GGGGGAAGATTTCAACCAATCGAAGAACCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 680
Qy 601 CCGGGGAGGCTCTGTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 681 CCGGGGAGGCTCTGTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740
Qy 661 AGGCGCACACACGCTTCAATGATTAATCCCAAGAGAGAGAGTCAATCGCTGCTGCTGCT 720
Db 741 AGGCGCACACACGCTTCAATGATTAATCCCAAGAGAGAGAGTCAATCGCTGCTGCTGCT 800
Qy 721 CGCTCAAGGCTTAATCTCAACACGCAAGGAGAGATGATGAGTGAAGAACTTAATC 780
Db 801 CGCTCAAGGCTTAATCTCAACACGCAAGGAGAGATGATGAGTGAAGAACTTAATC 860
Qy 781 CTACACAGAGACTACAGCGCTGACAGCTTGTCTACACACAGACATGCTGCTGCTGAG 840
Db 861 CTACACAGAGACTACAGCGCTGACAGCTTGTCTACACACAGACATGCTGCTGCTGAG 920
Qy 841 ATCCGTTCCAAAGAGAGGAGTGTGCGAGCCATCCCGGACTTACAGACATGCTGCTG 900
Db 921 ATCCGTTCCAAAGAGAGGAGTGTGCGAGCCATCCCGGACTTACAGACATGCTGCTGCT 980
Qy 901 CCCTGATGATTAATCAATCCCGAGTGTGCAACAGCTGATGATCACTGGCTTTGAGAAA 960
Db 981 CCCTGATGATTAATCAATCCCGAGTGTGCAACAGCTGATGATCACTGGCTTTGAGAAA 1040
Qy 961 GAGAAATTTACAGCAATATCTTAATCCGAGAGACAGCTGAAATGACTGTTGAGACTGAT 1020
Db 1041 GAGAAATTTACAGCAATATCTTAATCCGAGAGAGACAGCTGAAATGACTGTTGAGACTGAT 1100
Qy 1021 TCCACCGGAGAGTGTACAGAGCCCACTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1101 TCCACCGGAGAGTGTACAGAGCCCACTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160
Qy 1081 TGTGCTGTGACCCACAGTGTGAAACAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1161 TGTGCTGTGACCCACAGTGTGAAACAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1220
Qy 1221 GTCGTGCTCCCTCAAGGCGCATGACTTGAATGAGTGTGAGCTGGGCGCGTGATGT 1280
Qy 1201 GCCCTGAAGACAGAGCGGCTTACAGAGATTCACACTTTTACCTTGATGATCCG 1260
Db 1281 GCCCTGAAGACAGAGCGGCTTACAGAGATTCACACTTTTACCTTGATGATCCG 1340
Qy 1261 AGTCACACCAAGAGAGATGAGCTGGGCGCGTGATGT 1296
Db 1341 AGTCACACCAAGAGAGATGAGCTGGGCGCGTGATGT 1376
RESULT 11
US-09-221-107-122
; Sequence 122, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C2
; CURRENT APPLICATION NUMBER: US/09/221.107
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA

i ORGANISM: Human
US-09-221-107-122

Query Match 99.8%; Score 1292.8; DB 4; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGAGCCCTGCGCGCGCTGCTTCTCTGCGTCTGCTGCTGAGCGACTCCAAAGC 60
DB 81 ATGAGAGCCCTGCGCGCGCTGCTTCTCTGCGTCTGCTGCTGAGCGACTCCAAAGC 140
QY 61 AGCATGAACTTCAATCAAGTTCATCGAATCTGTAAGTCTTAAATGAGGAACTGTGTG 120
DB 141 AGCATGAACTTCAATCAAGTTCATCGAATCTGTAAGTCTTAAATGAGGAACTGTGTG 200
QY 121 TCCAACAAAGTACTTCTCCAACTTCACTGTGTGCAATGCGCCAAAGAAATTCGAGGCGAG 180
DB 201 TCCAACAAAGTACTTCTCCAACTTCACTGTGTGCAATGCGCCAAAGAAATTCGAGGCGAG 260
QY 181 CACTGTAAATAGATAGTCAAAAACCTGCTATAGAGGAAATGCTCACTTTTACCGAGA 240
DB 261 CACTGTAAATAGATAGTCAAAAACCTGCTATAGAGGAAATGCTCACTTTTACCGAGA 320
QY 241 AAGGCCAGCACTGACACCATGCGCGCGCTGCTGCTGCTGCACTGCGCACTGCTCT 300
DB 321 AAGGCCAGCACTGACACCATGCGCGCGCTGCTGCTGCTGCACTGCGCACTGCTCTCT 380
QY 301 CAGCAAAAGTACCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 381 CAGCAAAAGTACCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 440
QY 361 TACTGCAAGAACCCAGAACCCGAGAGCGAGCCCTGCTATGTCAGAGTGGCGCTAAAG 420
DB 441 TACTGCAAGAACCCAGAACCCGAGAGCGAGCCCTGCTATGTCAGAGTGGCGCTAAAG 500
QY 421 CCGCTTGTCCAAAGTGTGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 501 CCGCTTGTCCAAAGTGTGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
QY 481 CCAGAAAGATTTAAATTTCAAGTGTGCGCAAAAGCTCTGAGGCGCGCGCGCGCGCG 540
DB 561 CCAGAAAGATTTAAATTTCAAGTGTGCGCAAAAGCTCTGAGGCGCGCGCGCGCGCG 620
QY 541 GGGGAGAAATTCACCAATCGAAGAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 621 GGGGAGAAATTCACCAATCGAAGAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 680
QY 601 GGGGAGAAATTCACCAATCGAAGAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 681 GGGGAGAAATTCACCAATCGAAGAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 740
QY 661 AGGCCACACACTGCTTCAATTTAATCCAAAGAGAGGACTAATCGCTTCACTGAGG 720
DB 741 AGGCCACACACTGCTTCAATTTAATCCAAAGAGAGGACTAATCGCTTCACTGAGG 800
QY 721 CGCTCAAGGCTTAACTTCAACACGCAAGGGAGATGATGTTGAGGTGAGAAACCTAATC 780
DB 801 CGCTCAAGGCTTAACTTCAACACGCAAGGGAGATGATGTTGAGGTGAGAAACCTAATC 860
QY 781 CTACACAAAGACTACAGCGCTGACAGCTTGTGCAACCAAGCACTTGTGCTGAGG 840
DB 861 CTACACAAAGACTACAGCGCTGACAGCTTGTGCAACCAAGCACTTGTGCTGAGG 920
QY 841 ATCCGTTCCAAAGAGGAGAGGTGTGCGAGCCATCCCGGACTATACAGACATCTGCTG 900
DB 921 ATCCGTTCCAAAGAGGAGAGGTGTGCGAGCCATCCCGGACTATACAGACATCTGCTG 980
QY 901 CCGTCATATGTAATAGATCCCGAGTTTGGCAACAAGTGTGAGATCACTGCTTTGAGAAA 960
DB 981 CCGTCATATGTAATAGATCCCGAGTTTGGCAACAAGTGTGAGATCACTGCTTTGAGAAA 1040
QY 961 GAGAAATTTACAGCACTATCTTATCCGAGAGAGCTGAGAAATGATGTTGAGGCGATTT 1020

```

```

DB 1041 GAGAAATTTACAGCACTATCTTATCCGAGAGAGCTGAGAGATGATGTTGTGAGCTGATT 1100
QY 1021 TCCCAACGGGAGAGTGTGAGAGGCGCCCACTAATAGGCTCTGAAATGACCAACCAAAATGCTG 1080
DB 1101 TCCCAACGGGAGAGTGTGAGAGGCGCCCACTAATAGGCTCTGAAATGACCAACCAAAATGCTG 1160
QY 1081 TGTGCTGTGACCCACAGATGAGAAAAGATTTCTGCGCAGGAGAGCTCAGGGGAGCCCTC 1140
DB 1161 TGTGCTGTGACCCACAGATGAGAAAAGATTTCTGCGCAGGAGAGCTCAGGGGAGCCCTC 1220
QY 1141 GTCTGTTCCCTCCAAAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1221 GTCTGTTCCCTCCAAAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280
QY 1201 GCCCTGAAAGACAAAGCAGGCGCTTACACAGAGTCTCACCTTCTTACCTGATCCGC 1260
DB 1281 GCCCTGAAAGACAAAGCAGGCGCTTACACAGAGTCTCACCTTCTTACCTGATCCGC 1340
QY 1261 AGTCACACCAAGAAAGAAATGCGCTGCGCTCTGGA 1296
DB 1341 AGTCACACCAAGAAAGAAATGCGCTGCGCTCTGGA 1376

```

RESULT 12
518829-2
Patent No. 518829
APPLICANT: KOBAYASHI, YO-ICHI, OMORI, MUNEMI, YAMADA, CHIRAKO
TITLE OF INVENTION: RAPIDLY ACTING PROTEIN KINASE
NUMBER OF SEQUENCES: 23
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/340,007
FILING DATE: 18-AUG-1988
SEQ ID NO: 2
LENGTH: 2301

Query Match 99.8%; Score 1292.8; DB 6; Length 2301;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGAGCCCTGCGCGCGCTGCTTCTCTGCGTCTGCTGCTGAGCGACTCCAAAGC 60
DB 67 ATGAGAGCCCTGCGCGCGCTGCTTCTCTGCGTCTGCTGCTGAGCGACTCCAAAGC 126
QY 61 AGCAATGAATCTTCAATCAAGTTCATCGAATCTGTAAGTCTTAAATGAGGAACTGTGTG 120
DB 127 AGCAATGAATCTTCAATCAAGTTCATCGAATCTGTAAGTCTTAAATGAGGAACTGTGTG 186
QY 121 TCCAACAAAGTACTTCTCCAACTTCACTGTGTGCAATGCGCCAAAGAAATTCGAGGCGAG 180
DB 187 TCCAACAAAGTACTTCTCCAACTTCACTGTGTGCAATGCGCCAAAGAAATTCGAGGCGAG 246
QY 181 CACTGTAAATAGATAGTCAAAAACCTGCTATAGAGGAAATGCTCACTTTTACCGAGA 240
DB 247 CACTGTAAATAGATAGTCAAAAACCTGCTATAGAGGAAATGCTCACTTTTACCGAGA 306
QY 241 AAGGCCAGCACTGACACCAATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 307 AAGGCCAGCACTGACACCAATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
QY 301 CAGCAAAAGTACCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 367 CAGCAAAAGTACCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
QY 361 TACTGCAAGAACCCAGAACCCGAGAGCGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 427 TACTGCAAGAACCCAGAACCCGAGAGCGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 486
QY 421 CCGCTTGTCCAAAGAGTGTGATGATGCTGCGAGATGAGGAAAGAGCTCTCTCTCT 480
DB 487 CCGCTTGTCCAAAGAGTGTGATGATGCTGCGAGATGAGGAAAGAGCTCTCTCTCTCT 546
QY 481 CCAGAAAGATTTAAATTTCAAGTGTGCGCAAAAGACTCTGAGGCCCGCGCTTAAAGATTATT 540

```

Db 547 CCAGAGAAATTAAATTTGAGTGTGGCAAAAGACTGTAGGCCCCCTTTAGATTATT 606
Qy 541 GGGGGAGATTTCACCAATCGAAGAACAGCCCTGGTTTGGCCCATCTACAGAGGCAC 600
Db 607 GGGGGAGAAATTCACCAATCGAAGAACAGCCCTGGTTTGGCCCATCTACAGAGGCAC 666
Qy 601 GGGGGGGGCTGTGACCTACGTGTGTGGAGGAGAGCTTCAAGCCCTTGTGGGTATC 660
Db 667 GGGGGGGGCTGTGACCTACGTGTGTGGAGGAGAGCTTCAAGCCCTTGTGGGTATC 726
Qy 661 AGGCGACACACTGCTTCAATTGATTAACCAAGAGAGAGACTACATCGTCTACCTGGT 720
Db 727 AGGCGACACACTGCTTCAATTGATTAACCAAGAGAGAGACTACATCGTCTACCTGGT 786
Qy 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAAGTTGAGGTGAAAACTTAATC 780
Db 787 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAAGTTGAGGTGAAAACTTAATC 846
Qy 781 CTACACAAGGACTACAGCGCTGACAGCTTGTGCTACCAACAAGCAATGTGCTTGTGAAG 840
Db 847 CTACACAAGGACTACAGCGCTGACAGCTTGTGCTACCAACAAGCAATGTGCTTGTGAAG 906
Qy 841 ATCCGTTCCAAAGAGGAGGAGGTGTGCGACAGCCATCCCGGACTATACAGACCATCTGCTG 900
Db 907 ATCCGTTCCAAAGAGGAGGAGGTGTGCGACAGCCATCCCGGACTATACAGACCATCTGCTG 966
Qy 901 CCTCGATGTATACGATCCCAAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGAAAA 960
Db 967 CCTCGATGTATACGATCCCAAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGAAAA 1026
Qy 961 GAGAAATTCACCAATCTCTATTCGGAGACACTGAAAAATGACTGTGTGAAGCTGATT 1020
Db 1027 GAGAAATTCACCAATCTCTATTCGGAGACACTGAAAAATGACTGTGTGAAGCTGATT 1086
Qy 1021 TCCCAACGGGAGTGTAGCAGCCCACTACTACGAGCTGTGAATGCAACCAAAATGCTG 1080
Db 1087 TCCCAACGGGAGTGTAGCAGCCCACTACTACGAGCTGTGAATGCAACCAAAATGCTG 1146
Qy 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGTGCAAGGAGACTCAAGGGAGACCCCTC 1140
Db 1147 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGTGCAAGGAGACTCAAGGGAGACCCCTC 1206
Qy 1141 GTCTGTTCCTCCAAAGCCGCACTTGTGACTGGAATGTGTAGCTGGGGCCGTGATGT 1200
Db 1207 GTCTGTTCCTCCAAAGCCGCACTTGTGACTGGAATGTGTAGCTGGGGCCGTGATGT 1266
Qy 1201 GGCCTGAAGACAAAGCCAGGCGCTCAACAGAGAGTCAACACTTCTTAACCTGGATCCGC 1260
Db 1267 GGCCTGAAGACAAAGCCAGGCGCTCAACAGAGAGTCAACACTTCTTAACCTGGATCCGC 1326
Qy 1261 AGTCACACCAAGAGAGAGATGGCTGGCCCTTGA 1296
Db 1327 AGTCACACCAAGAGAGAGATGGCTGGCCCTTGA 1362

RESULT 13
5219569-1
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO:1:
; LENGTH: 1372
5219569-1

Query Match 99.4%; Score 1288; DB 6; Length 1372;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ATGAGAGCCCTGTGCGGCGCCCTGCTTCTCTGCTCTGCTGCTGAGCGACTCAAGGC 60
Db 77 ATGAGAGCCCTGTGCGGCGCCCTGCTTCTCTGCTCTGCTGCTGAGCGACTCAAGGC 136
Qy 61 AGCAATGAACCTTCATCAAGTTCATGAACTGTGACTGTCTTAATGAGAGACATGTGTG 120
Db 137 AGCAATGAACCTTCATCAAGTTCATGAACTGTGACTGTCTTAATGAGAGACATGTGTG 196
Qy 121 TCCAAAGATCTTCTCAACATTCGCTGTGCACTGCGCCCAAGAAATTCGAGGGCAG 180
Db 197 TCCAAAGATCTTCTCAACATTCGCTGTGCACTGCGCCCAAGAAATTCGAGGGCAG 256
Qy 181 CACTGTGAATTAAGTAAAGTCAAAACCTGTATGAGGGGAAATGCTACTTTTACGAGGA 240
Db 257 CACTGTGAATTAAGTAAAGTCAAAACCTGTATGAGGGGAAATGCTACTTTTACGAGGA 316
Qy 241 AAGGCCAGCACTGACACATGAGCGCCCTGCTGCTGCTGCAACTGTGCACTGTCTT 300
Db 317 AAGGCCAGCACTGACACATGAGCGCCCTGCTGCTGCTGCAACTGTGCACTGTCTT 376
Qy 301 CAGCAACGTAACATGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 377 CAGCAACGTAACATGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 436
Qy 361 TACTGCGAGAACCCAGACAAACCGGAGCGACCCCTGTGCTATGTGAGGTGAGCTTAAG 420
Db 437 TACTGCGAGAACCCAGACAAACCGGAGCGACCCCTGTGCTATGTGAGGTGAGCTTAAG 496
Qy 421 CGGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 497 CGGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556
Qy 481 CCAGAGAAATTAATTAAGTGTGCGCAAAAGACTGTAGGCGCCGCTTTAAGATTATT 540
Db 557 CCAGAGAAATTAATTAAGTGTGCGCAAAAGACTGTAGGCGCCGCTTTAAGATTATT 616
Qy 541 GGGGGAGAAATTCACCAATCGAAGAACAGCCCTGGTTTGGCCCATCTACAGAGGCAC 600
Db 617 GGGGGAGAAATTCACCAATCGAAGAACAGCCCTGGTTTGGCCCATCTACAGAGGCAC 676
Qy 601 CGGGGGGGCTGTGTCACTTGTGTGAGGAGGAGCTATACAGCCCTTGTGCTGCTGATC 660
Db 677 CGGGGGGGCTGTGTCACTTGTGTGAGGAGGAGCTATACAGCCCTTGTGCTGCTGATC 736
Qy 661 AGGCGCACACTGCTTCAATTGATTAACCAAGAGAGAGACTACATCGTCTTACCTGGGT 720
Db 737 AGGCGCACACTGCTTCAATTGATTAACCAAGAGAGAGACTACATCGTCTTACCTGGGT 796
Qy 721 CGCTCAAGGCTTAATCTCAACAGCAAGGGAGATGAAGTTGAGGTGAAAACTTAATC 780
Db 797 CGCTCAAGGCTTAATCTCAACAGCAAGGGAGATGAAGTTGAGGTGAAAACTTAATC 856
Qy 781 CTACACAAGGACTACAGCGCTGACAGCTTGTGCTACCAACAAGCAATGTGCTTGTGAAG 840
Db 857 CTACACAAGGACTACAGCGCTGACAGCTTGTGCTACCAACAAGCAATGTGCTTGTGAAG 916
Qy 841 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCCGGACTTAACAGCAATGTGCTG 900
Db 917 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCCGGACTTAACAGCAATGTGCTG 976
Qy 901 CCTCGATGTATACGATCCCAAGTTTGGCAACAAGTGTGAGATCACTGGCTTTGAAAA 960
Db 977 CCTCGATGTATACGATCCCAAGTTTGGCAACAAGTGTGAGATCACTGGCTTTGAAAA 1036
Qy 961 GAGAAATTCACCAATCTCTATTCGGAGACACTGAAAAATGACTGTGTGAGACTGATT 1020
Db 1037 GAGAAATTCACCAATCTCTATTCGGAGACACTGAAAAATGACTGTGTGAGACTGATT 1096
Qy 1021 TCCCAACGGGAGTGTAGCAGCCCACTACTACGAGCTGTGAAGTCAACCAAAATGCTG 1080

Db 1097 TCCACCGGAGTGTGAGAGCCCACTAAGCTGTGAAGTCAACCAATATGCTG 1156
Qy 1081 TGTGCTGTGACCAAGTGTGAAAACAGATTCTGTGCGAGGAGACTAGGGGAGCCCTC 1140
Db 1157 TGTGCTGTGACCAAGTGTGAAAACAGATTCTGTGCGAGGAGACTAGGGGAGCCCTC 1216
Qy 1141 GTCTGTTTCCCTCCAAAGGCGGCAATGCTTTGACTGTGAATTGTAGCTGGGGCCGTGATGT 1200
Db 1217 GTCTGTTTCCCTCCAAAGGCGGCAATGCTTTGACTGTGAATTGTAGCTGGGGCCGTGATGT 1276
Qy 1201 GCCCTGAAGACAAGCGGAGCTGTACAGAGAGTGTACACTTAACTTGAATCCG 1260
Db 1277 GCCCTGAAGACAAGCGGAGCTGTACAGAGAGTGTACACTTAACTTGAATCCG 1336
Qy 1261 AGTCACACCAAGAAAGAGATGGCTGTGCTGTGA 1296
Db 1337 AGTCACACCAAGAAAGAGATGGCTGTGCTGTGA 1372

RESULT 14

US-07-957-039A-7
; Sequence 7, Application US/07957039A
; Patent No. 5389538
; GENERAL INFORMATION:
; APPLICANT: TANABE, TOSHIKAZUMI
; APPLICANT: MORITA, MASANORI
; APPLICANT: HIROSE, MASAKI
; APPLICANT: AMASUJI, YASUO
; TITLE OF INVENTION: MUTANT HUMAN PROHOKINASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/957,039A
; FILING DATE: 06-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 289257/1991
; FILING DATE: 07-OCT-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1233
US-07-957-039A-7

Query Match 95.1%; Score 1232.8; DB 1; Length 1236;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 61 AGCAATGAACTTCAATCAAGTTCATGCAACTGTGACTGTCTAATAATGAGAGAACTGTGTG 120

Db 1 AGCAATGAACTTCAATCAAGTTCATGCAACTGTGACTGTCTAATAATGAGAGAACTGTGTG 60
Qy 121 TCCAAAGTACTTCTCCCAACATTCATGTGTGCACTGTGCAAGAAATTTGGAGGGCAG 180
Db 61 TCCAAAGTACTTCTCCCAACATTCATGTGTGCACTGTGCAAGAAATTTGGAGGGCAG 120
Qy 181 CACTGTGAATATGATTAAGTCAAAAACCTGTCTATGAGGGAAATGTCACTTTTACGAGGA 240
Db 121 CACTGTGAATATGATTAAGTCAAAAACCTGTCTATGAGGGAAATGTCACTTTTACGAGGA 180
Qy 241 AAGGCGAGACTGACCAATGTGGCCGCGCTGTGCTGCTGCTGAACTGTGCACTGTCTT 300
Db 181 AAGGCGAGACTGACCAATGTGGCCGCGCTGTGCTGCTGCTGAACTGTGCACTGTCTT 240
Qy 301 CAGCAAAAGTACATGCGCAAGATGTGATGTCTTCAAGTGGGCTGGGGAAACATAT 360
Db 241 CAGCAAAAGTACATGCGCAAGATGTGATGTCTTCAAGTGGGCTGGGGAAACATAT 300
Qy 361 TACTCAGGAAACCAAGACACCGAGGCGACCTGTGTCTATGTGCAAGTGGGCTTAAAG 420
Db 301 TACTCAGGAAACCAAGACACCGAGGCGACCTGTGTCTATGTGCAAGTGGGCTTAAAG 360
Qy 421 CGGCTGTCCAAAGATGATGTGTGATGATGCGGCAATGGAATAAAGCCCTCTCTCT 480
Db 361 CGGCTGTCCAAAGATGATGTGTGATGATGCGGCAATGGAATAAAGCCCTCTCTCT 420
Qy 481 CCAGAAAGATTAATTAATTTCAATGTGCGCAAAAGACTGTGAGGCGCCCTTTAATTA 540
Db 421 CCAGAAAGATTAATTAATTTCAATGTGCGCAAAAGACTGTGAGGCGCCCTTTAATTA 480
Qy 541 GGGGGAATTCACCAATGTGAGAACCGACCTGTGTTGGGCACTTACAGAGGCAAC 600
Db 481 GGGGGAATTCACCAATGTGAGAACCGACCTGTGTTGGGCACTTACAGAGGCAAC 540
Qy 601 CGGGGGGCTGTGACCTAAGTGTGAGGAGGCGCTCAATGACCTTGTGTGTATC 660
Db 541 CGGGGGGCTGTGACCTAAGTGTGAGGAGGCGCTCAATGACCTTGTGTGTATC 600
Qy 661 AGCGGCAACA CTGCTTCAATGATTAACCAAGAGAGAGACTATCATGTCTACCTGGGT 720
Db 601 AGCGGCAACA CTGCTTCAATGATTAACCAAGAGAGAGACTATCATGTCTACCTGGGT 660
Qy 721 CGCTCAAGGCTTAATCTCAAAAGGAGAGATGAAGTTGAGGTGAGAAACTTATC 780
Db 661 CGCTCAAGGCTTAATCTCAAAAGGAGAGATGAAGTTGAGGTGAGAAACTTATC 720
Qy 781 CTACACAAGGACTACAGGCTGACAGGCTTGTCAACCAACAGCATTTGCTGTGAG 840
Db 721 CTACACAAGGACTACAGGCTGACAGGCTTGTCAACCAACAGCATTTGCTGTGAG 780
Qy 841 ATCGGTTCCAAAGAGGAGGTGTGCGCAAGCATCCGGGACTATACAGCATCTGCTG 900
Db 781 ATCGGTTCCAAAGAGGAGGTGTGCGCAAGCATCCGGGACTATACAGCATCTGCTG 840
Qy 901 CCTCGATGATTAACGATTCCTCCAGTTTGGCAAGCTGTGATCACTGTGCTTTGAGAA 960
Db 841 CCTCGATGATTAACGATTCCTCCAGTTTGGCAAGCTGTGATCACTGTGCTTTGAGAA 900
Qy 961 GAGAAATTCACGACTATCTCTATCCGAGACAGCTGAAAATGACTGTGGAAGTGAATT 1020
Db 901 GAGAAATTCACGACTATCTCTATCCGAGACAGCTGAAAATGACTGTGGAAGTGAATT 960
Qy 1021 TCCCAACGGGAGTGTGACAGCCCACTACTACAGGCTGTGAATGTCAACCAAAATGTG 1080
Db 961 TCCCAACGGGAGTGTGACAGCCCACTACTACAGGCTGTGAATGTCAACCAAAATGTG 1020
Qy 1081 TGTGCTGTGACCAAGTGTGAAAACAGATTCTGTGCGAGGAGACTCAAGGGGAGCCCTC 1140
Db 1021 TGTGCTGTGACCAAGTGTGAAAACAGATTCTGTGCGAGGAGACTCAAGGGGAGCCCTC 1080
Qy 1141 GTCTGTTTCCCTCCAAAGGCGGCAATGCTTTGACTGTGAATTGTGAGCTGGGGCCGTGATGT 1200

Db 1081 GTCGTCTCCCTCCAGGCGCATGATCTTGTGATGATGAGCTGGGCGGTGATGT 1140
QY 1201 GGCCTAAGAGCAAGCAGGCGCTCTACAGAGTCTCACTTTTACCTTGAGATCCGC 1260
Db 1141 GCCCTAAGAGCAAGCAGGCGCTCTACAGAGTCTCACTTTTACCTTGAGATCCGC 1200
QY 1261 AGTCACCAAGAGAGAGATGAGCTGGCCCTCTGA 1296
Db 1201 AGTCACCAAGAGAGAGATGAGCTGGCCCTCTGA 1236

RESULT 15
US-09-023-655-927
; Sequence 927, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 927:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 91311467
; US-09-023-655-927

Query Match 95.1%; Score 1232.8; DB 4; Length 1236;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 61 AGCAATGAACTTCATCAAGTTCATGCACTGCTGTCTAAATGAGAAATGTGTG 120
Db 1 AGCAATGAACTTCATCAAGTTCATGCACTGCTGTCTAAATGAGAAATGTGTG 60
QY 121 TCACAACAGTACTTCTCCCAACTTCACTGTGCAACTGCCCAAGAAATTCGAGGCGAG 180
Db 61 TCACAACAGTACTTCTCCCAACTTCACTGTGCAACTGCCCAAGAAATTCGAGGCGAG 120
QY 181 CACTGGAATATGATTAAGTCAAAAACCTGCTATGAGGGAGATGCTACTTTTACGAGAGA 240

Db 121 CACTGGAATATGATTAAGTCAAAAACCTGCTATGAGGGAGATGCTACTTTTACGAGAGA 180
QY 241 AAGGCGAGACTGACACACATGAGCGCGGCTGCTGCTCCCTGGAATCTGCACTGTCTT 300
Db 181 AAGGCGAGACTGACACACATGAGCGCGGCTGCTGCTCCCTGGAATCTGCACTGTCTT 240
QY 301 CAGCAACGTAACCATGCCCCACAGTGTGCTTCACTGGGCTGGGGAAACATTAAT 360
Db 241 CAGCAACGTAACCATGCCCCACAGTGTGCTTCACTGGGCTGGGGAAACATTAAT 300
QY 361 TACTGAGAGAACCCACAGACACGAGCGGACCCCTGCTGCTATGTGAGGTGAGGCTTAAG 420
Db 301 TACTGAGAGAACCCACAGACACGAGCGGACCCCTGCTGCTATGTGAGGTGAGGCTTAAG 360
QY 421 CCGCTTTCOAAGATGCACTGCTGCACTGCTGCACTGCTGCACTGCTGCTGCTGCT 480
Db 361 CCGCTTTCOAAGATGCACTGCTGCACTGCTGCACTGCTGCACTGCTGCTGCTGCTGCT 420
QY 481 CCAGAAATTTAAATTTAAGTGTGCGCAAAAGACTGTAGAGCCGCTTTAAGATTAAT 540
Db 421 CCAGAAATTTAAATTTAAGTGTGCGCAAAAGACTGTAGAGCCGCTTTAAGATTAAT 480
QY 541 GGGGAGAAATTCACACATGCAAGAGACAGCCCTGCTTGGGCTTACAGAGAGCAC 600
Db 481 GGGGAGAAATTCACACATGCAAGAGACAGCCCTGCTTGGGCTTACAGAGAGCAC 540
QY 601 CGGGGGGGCTGTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 541 CGGGGGGGCTGTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 661 AAGGCGAGACTGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACT 720
Db 601 AAGGCGAGACTGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACT 660
QY 721 CGCTCAAGGCTTAACTCCCAACGCAAGGGAGATGAGTTGAGGTGAGAAACCTAATC 780
Db 661 CGCTCAAGGCTTAACTCCCAACGCAAGGGAGATGAGTTGAGGTGAGAAACCTAATC 720
QY 781 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACAGACATTCCTGCTGAAG 840
Db 721 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACAGACATTCCTGCTGTAAG 780
QY 841 ATCCGTTCCAAAGAGGCGAGGTGTGCGGACCATCCCGGACTATACAGACATTCCTG 900
Db 781 ATCCGTTCCAAAGAGGCGAGGTGTGCGGACCATCCCGGACTATACAGACATTCCTG 840
QY 901 CCCTGATGTATTAAGATCCCGAGTTTGGACAAAGCTGTGAGATCACTGCTTGGAGAA 960
Db 841 CCCTGATGTATTAAGATCCCGAGTTTGGACAAAGCTGTGAGATCACTGCTTGGAGAA 900
QY 961 GAGAAATTCACGACTATCTATCCGAGAGAGCTGAGAAATGATCTGTTGAGAGCTGATT 1020
Db 901 GAGAAATTCACGACTATCTATCCGAGAGAGCTGAGAAATGATCTGTTGAGAGCTGATT 960
QY 1021 TCCCAACCGGAGATGTGAGAGCCCACTATACGCTCTGAAAGTCAACCAAAATGCTG 1080
Db 961 TCCCAACCGGAGATGTGAGAGCCCACTATACGCTCTGAAAGTCAACCAAAATGCTG 1020
QY 1081 TGTGCTGCTGACCCACAGTGTGAGAAACAGATTTCTGCAAGGAGACTCAAGGGGAGCCCTTC 1140
Db 1021 TGTGCTGCTGACCCACAGTGTGAGAAACAGATTTCTGCAAGGAGACTCAAGGGGAGCCCTTC 1080
QY 1141 GTCGTCTCCCTCCAGGCGGCACTGACTTGAATGAGAAATTTGAGAGTGGGCGGTGAGATGT 1200
Db 1081 GTCGTCTCCCTCCAGGCGGCACTGACTTGAATGAGAAATTTGAGAGTGGGCGGTGAGATGT 1140
QY 1201 GGCCTGAAGAGACAGGCGGCTTACACAGAGTCTCACTTTTACCTTGATCCGC 1260
Db 1141 GGCCTGAAGAGACAGGCGGCTTACACAGAGTCTCACTTTTACCTTGATCCGC 1200
QY 1261 AGTCACCAAGAGAGAGATGAGCTGGCCCTCTGA 1296

Db 1201 AGTCACCAAGAGAGATGGCTGGCCCTTGA 1236

Search completed: March 21, 2004, 20:52:09
Job time : 128 secs

121 TCACAAGATGTTCTCCAAAGATTGACGTGTGCAATCTGCCAAGAAATTTGGAGGGCAG 180
121 TCACAAGATGTTCTCCAAAGATTGACGTGTGCAATCTGCCAAGAAATTTGGAGGGCAG 180
181 CACTGTGAATAGATTAAGTCAAAAACTGTATAGAGGGAATGGTCACTTTTACCGAGGA 240
181 CACTGTGAATAGATTAAGTCAAAAACTGTATAGAGGGAATGGTCACTTTTACCGAGGA 240
241 AAGGCGACGACTGACACCAATGGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
241 AAGGCGACGACTGACACCAATGGGCGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 300
301 CAGCAAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
301 CAGCAAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
361 TACTGACAGAAACCCAGAACCCGAGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 420
361 TACTGACAGAAACCCAGAACCCGAGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 420
421 CCGCTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
421 CCGCTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
481 CCAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
481 CCAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
541 GGGGGAAGATTCACCAACCATGAGAACCAAGCCCTGTGCTGCTGCTGCTGCTGCTGCTG 600
541 GGGGGAAGATTCACCAACCATGAGAACCAAGCCCTGTGCTGCTGCTGCTGCTGCTGCTG 600
601 CCGGGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
601 CCGGGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
661 AGCCGCAACACTGCTTCAATGATTACCCAAAGAGAGAGACTACATGCTCTGCTGCTG 720
661 AGCCGCAACACTGCTTCAATGATTACCCAAAGAGAGAGACTACATGCTCTGCTGCTGCT 720
721 CCGCTCAAGGCTTAACTCCAAACGCAAGGGGAGATGAAATTTGAGGGAATCTTAACT 780
721 CCGCTCAAGGCTTAACTCCAAACGCAAGGGGAGATGAAATTTGAGGGAATCTTAACT 780
781 CTACACAAGAGACTACAGGCTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
781 CTACACAAGAGACTACAGGCTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
841 ATCCGCTTCCAAAGAGGCGAGGTGTGCGACGCAATCCCGGACTATACAGACCATCTG 900
841 ATCCGCTTCCAAAGAGGCGAGGTGTGCGACGCAATCCCGGACTATACAGACCATCTG 900
901 CCGCTGATGATTAACGATTCGCCAGTTTGGCAAGAGCTGAGATTCATGCTGCTTGGAAA 960
901 CCGCTGATGATTAACGATTCGCCAGTTTGGCAAGAGCTGAGATTCATGCTGCTTGGAAA 960
961 GAGAAATTCACGACTATCTATCCGAGACAGCTGAAAATGACTGTTGGAAGCTGATT 1020
961 GAGAAATTCACGACTATCTATCCGAGACAGCTGAAAATGACTGTTGGAAGCTGATT 1020
1021 TCCCAACCGGAGAGTGTGACGAGCCCACTACTAGGCTCTGAAATGACCAACCAATGCTG 1080
1021 TCCCAACCGGAGAGTGTGACGAGCCCACTACTAGGCTCTGAAATGACCAACCAATGCTG 1080
1081 TGTGCTGTGACCCACAGTGTGAAAGAGATTCCTGTGCAAGAGAGACTCAAGGAGAACCC 1140
1081 TGTGCTGTGACCCACAGTGTGAAAGAGATTCCTGTGCAAGAGAGACTCAAGGAGAACCC 1140
1141 GTCTGTTCCTTCCAAAGCGGAGTGAATTTGATGAAATTTGAGTGTGAGGCGCTGGAAT 1200
1141 GTCTGTTCCTTCCAAAGCGGAGTGAATTTGATGAAATTTGAGTGTGAGGCGCTGGAAT 1200
1201 GCCCTGAAGAGCAAGCGGCGCTGTCACAGAGGTCTGACACTTCTTACCTGTGATCCGC 1260

Db 1201 GCCCTGAAGAGCAAGCGGCGCTGTCACAGAGGTCTGACACTTCTTACCTGTGATCCGC 1260
Qy 1261 AGTCACACCAAGAGAGAGATGGCCCTGAGCTTGA 1296
Db 1261 AGTCACACCAAGAGAGAGATGGCCCTGAGCTTGA 1296
RESULT 2
US-09-735-705-123
Sequence 123, Application US/09735705
Patent No. US20020052329A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aljun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C14
CURRENT APPLICATION NUMBER: US/09/735,705
CURRENT FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 419
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 123
LENGTH: 2294
TYPE: DNA
ORGANISM: Homo sapien
US-09-735-705-123
Query Match 100.0%; Score 1296; DB 9; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAGCGGCTGAGGCGGCGCTGCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 67 ATGAGAGCGGCTGAGGCGGCGCTGCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
Qy 61 AGCAATGAATCTTCAATCAAGTTCATGCACTGCTGCACTGCTGCTGCTGCTGCTGCTGCT 120
Db 127 AGCAATGAATCTTCAATCAAGTTCATGCACTGCTGCACTGCTGCTGCTGCTGCTGCTGCT 186
Qy 121 TCACAAGATGTTCTCCAAAGATTGACGTGTGCAATCTGCCAAGAAATTTGGAGGGCAG 180
Db 187 TCACAAGATGTTCTCCAAAGATTGACGTGTGCAATCTGCCAAGAAATTTGGAGGGCAG 246
Qy 181 CACTGTGAATAGATTAAGTCAAAAACTGTATAGAGGGAATGGTCACTTTTACCGAGGA 240
Db 247 CACTGTGAATAGATTAAGTCAAAAACTGTATAGAGGGAATGGTCACTTTTACCGAGGA 306
Qy 241 AAGGCGACGACTGACACCAATGGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 307 AAGGCGACGACTGACACCAATGGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
Qy 301 CAGCAAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
Db 367 CAGCAAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 361 TACTGACAGAAACCCAGAACCCGAGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 427 TACTGACAGAAACCCAGAACCCGAGGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 486
Qy 421 CCGCTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 487 CCGCTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546

```
QY 481 CCAGAGAAATTAATTTGAGTGGCCAAAGACTGTAGGGCCCGCTTTAGATTATT 540
Db 547 CCAAGAAATTAATTAATTTGAGTGGCCAAAGACTGTAGGGCCCGCTTTAGATTATT 606
QY 541 GGGGGAGAAATTCACCAATCGAAGAACAGCCCTGGTTTGGCGCATCTTACAGAGGAC 600
Db 607 GGGGGAGAAATTCACCAATCGAAGAACAGCCCTGGTTTGGCGCATCTTACAGAGGAC 666
QY 601 GGGGGGGGCTGTCTACCTTACCTGTGTGAGAGGCAAGCTTACAGCCCTTGTGGTATC 660
Db 667 GGGGGGGGCTGTCTACCTTACCTGTGTGAGAGGCAAGCTTACAGCCCTTGTGGTATC 726
QY 661 AGGGCCACACACTGCTTCAATTGATTAACCAAGAGAGGACTTACCTGCTACCTGGGT 720
Db 727 AGGGCCACACACTGCTTCAATTGATTAACCAAGAGAGGACTTACCTGCTACCTGGGT 786
QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGAATTGAGGTGAGAAAACCTAATC 780
Db 787 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGAATTGAGGTGAGAAAACCTAATC 846
QY 781 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACACAGCATTTGCTTGTGAAG 840
Db 847 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACACAGCATTTGCTTGTGAAG 906
QY 841 ATCCGTTCCAAAGAGGAGGAGTGTGGGAGCCATCCCGGACTTATACAGACATCTGCTG 900
Db 907 ATCCGTTCCAAAGAGGAGGAGTGTGGGAGCCATCCCGGACTTATACAGACATCTGCTG 966
QY 901 CCTCGATGTATTAACGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGGAAA 960
Db 967 CCTCGATGTATTAACGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGGAAA 1026
QY 961 GGAATTTCTACCGACTATCTTATCCGAGACAGCTGAGAAATGACTGTGTGAAGTATT 1020
Db 1027 GGAATTTCTACCGACTATCTTATCCGAGAGCAAGCTGAGAAATGACTGTGTGAAGTATT 1086
QY 1021 TCCCAACCGGAGATGTACAGAGCCCACTTACAGGCTGTGAAGTCCACCAAAAGCTG 1080
Db 1087 TCCCAACCGGAGATGTACAGAGCCCACTTACAGGCTGTGAAGTCCACCAAAAGCTG 1146
QY 1081 TGTGCTGTGACCCACAGATGAGAAAACAGATTCTGTGCAGAGAGACTCAGGGGAGACCCCTC 1140
Db 1147 TGTGCTGTGACCCACAGATGAGAAAACAGATTCTGTGCAGAGAGACTCAGGGGAGACCCCTC 1206
QY 1141 GTCTGTTCCTTCAGAGCCGATGACTTGAATGTGAGCTGGGGCCGTGTGATGT 1200
Db 1207 GTCTGTTCCTTCAGAGCCGATGACTTGAATGTGAGCTGGGGCCGTGTGATGT 1266
QY 1201 GGGCTGAAGAGACAGCCAGGCGCTTACAGAGAGTCTCACTTTACCTGTGATCCGC 1260
Db 1267 GGGCTGAAGAGACAGCCAGGCGCTTACAGAGAGTCTCACTTTACCTGTGATCCGC 1326
QY 1261 AGTCACACCAAGAGAGAGATGGCCCTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAGATGGCCCTGGCCCTCTGA 1362
```

```
RESULT 3
US-09-850-716A-123
; Sequence 123. Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Reiter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: Fastseq for Windows Version 3.0
```

```
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-850-716A-123
Query Match 100.0%; Score 1296; DB 9; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAGCCCTGCTGGCGGCGCTGTCTTCTGTGGTCTGTGAGGACTTCGCAAGGC 60
Db 67 ATGAGAGCCCTGCTGGCGGCGCTGTCTTCTGTGGTCTGTGAGGACTTCGCAAGGC 126
QY 61 AGCAATGAATCTTCAATGAATGTTGAACTGTGATGTCTTAATGAGAGACATGTGTG 120
Db 127 AGCAATGAATCTTCAATGAATGTTGAACTGTGATGTCTTAATGAGAGACATGTGTG 186
QY 121 TCCAAAGACTACTCTCCCAACATTCAGTGTGCACTGCCCCAAGAAATTCGGAGGCGAG 180
Db 187 TCCAAAGACTACTCTCCCAACATTCAGTGTGCACTGCCCCAAGAAATTCGGAGGCGAG 246
QY 181 CACTGGAATAGATTAAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTTACCGAGA 240
Db 247 CACTGGAATAGATTAAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTTACCGAGA 306
QY 241 AAGGCCAGACTGACACCATGGGCGGCGCTGCTGCTGCGAATCTTGCACTGTCTT 300
Db 307 AAGGCCAGACTGACACCATGGGCGGCGCTGCTGCTGCGAATCTTGCACTGTCTT 366
QY 301 CAGCAAGCTACATGCCCCACAGATCTGATGTCTTCACTGGGCGCTGGGGAAACATAAT 360
Db 367 CAGCAAGCTACATGCCCCACAGATCTGATGTCTTCACTGGGCGCTGGGGAAACATAAT 426
QY 361 TACTGAGAGAACCCAGACACACCGAGGCGACCTGTGTCTATGTGAGGTGGGCTTAAG 420
Db 427 TACTGAGAGAACCCAGACACACCGAGGCGACCTGTGTCTATGTGAGGTGGGCTTAAG 486
QY 421 CCGCTTGTCAAGATGATGTGTGATGATCTGCGCAGATGAGAAAAGCCCTCTCTCT 480
Db 487 CCGCTTGTCAAGATGATGTGTGATGATCTGCGCAGATGAGAAAAGCCCTCTCTCT 546
QY 481 CCAGAGAAATTAATTTCAAGTGTGGCCAAAGACTGTAGGCCCCGCTTTAAGATTATT 540
Db 547 CCAGAGAAATTAATTTCAAGTGTGGCCAAAGACTGTAGGCCCCGCTTTAAGATTATT 606
QY 541 GGGGGAGAAATTCACCAATCGAAGAACAGCCCTGGTTTGGCGCATCTTACAGAGGAC 600
Db 607 GGGGGAGAAATTCACCAATCGAAGAACAGCCCTGGTTTGGCGCATCTTACAGAGGAC 666
QY 601 GGGGGGGGCTGTCTACCTTACCTGTGTGAGAGGCAAGCTTACAGCCCTTGTGGTATC 660
Db 667 GGGGGGGGCTGTCTACCTTACCTGTGTGAGAGGCAAGCTTACAGCCCTTGTGGTATC 726
QY 661 AGGGCCACACACTGCTTCAATTGATTAACCAAGAGAGGACTTACCTGCTACCTGGGT 720
Db 727 AGGGCCACACACTGCTTCAATTGATTAACCAAGAGAGGACTTACCTGCTACCTGGGT 786
QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGAATTGAGGTGAGAAAACCTAATC 780
Db 787 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGAATTGAGGTGAGAAAACCTAATC 846
QY 781 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACACAGCATTTGCTTGTGAAG 840
Db 847 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACACAGCATTTGCTTGTGAAG 906
QY 841 ATCCGTTCCAAAGAGGAGGAGTGTGGGAGCCATCCCGGACTTATACAGACATCTGCCC 900
Db 907 ATCCGTTCCAAAGAGGAGGAGTGTGGGAGCCATCCCGGACTTATACAGACATCTGCCC 966
QY 901 CCTCGATGTATTAACGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGGAAA 960
Db 967 CCTCGATGTATTAACGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGGAAA 1026
```

QY 961 GAGATTCTACCGACTATCTATCCGAGCAGCTGAAAATGACTGTGTGAACCTGATT 1020
|||
Db 1027 GAGAAATTCACGACTATCTATCTCCGAGCAGCTGAAAATGACTGTGTGAACCTGATT 1086
|||
QY 1021 TCCCAACCGGAGTGTCAAGCCCACTACTACGGCTGTGAAGTCAACCAAAATGCTG 1080
|||
Db 1087 TCCCAACCGGAGTGTCAAGCCCACTACTACGGCTGTGAAGTCAACCAAAATGCTG 1146
|||
QY 1081 TGTGTCTGTGACCCCAAGTGAAGAAACAGATTCTTCCCAAGGAGACTAGAGGGAGCCCTC 1140
|||
Db 1147 TGTGTCTGTGACCCCAAGTGAAGAAACAGATTCTTCCCAAGGAGACTAGAGGGAGCCCTC 1206
|||
QY 1141 GTCTGTTCCCTCCAGGCGGATGACTTTGAGTGTGAGCTGGGGCCGTGATGT 1200
|||
Db 1207 GTCTGTTCCCTCCAGGCGGATGACTTTGAGTGTGAGCTGGGGCCGTGATGT 1266
|||
QY 1201 GCCCTGAAGACAAAGCCAGGCGTCTACACAGAGTCTACACTTTACCTGGATCCGC 1260
|||
Db 1267 GCCCTGAAGACAAAGCCAGGCGTCTACACAGAGTCTACACTTTACCTGGATCCGC 1326
|||
QY 1261 AGTACACCAAGAAAGAAATGGCTGGCCCTCTGA 1296
|||
Db 1327 AGTACACCAAGAAAGAAATGGCTGGCCCTCTGA 1362
|||

RESULT 4

US-09-897-778-123
; Sequence 123. Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-123

Query Match 100.0%; Score 1296; DB 9; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGTGCGCGGCTGTCTCTGTGCTCTGTGCTGAGCGACTCCAAAGGC 60
|||
Db 67 ATGAGAGCCCTGTGCGCGGCTGTCTCTGTGCTCTGTGCTGAGCGACTCCAAAGGC 126
|||
QY 61 AGCAATGAAGCTTCATCAAGTTCATCGAAGCTGTGACTGTCTAAATGGAGAAACATGTGTG 120
|||
Db 127 AGCAATGAAGCTTCATCAAGTTCATCGAAGCTGTGACTGTCTAAATGGAGAAACATGTGTG 186
|||
QY 121 TCCAAAGTACTTCTCAACATTCATCTGTGCACTGCGCCAAAGAAATGGAGGGGAG 180
|||
Db 187 TCCAAAGTACTTCTCAACATTCATCTGTGCACTGCGCCAAAGAAATGGAGGGGAG 246
|||
QY 181 CACTGTGAATAGATAAGTCAAAAACCTGTCTATGAGGGGAATGATCACTTTTACCGAGGA 240
|||
Db 247 CACTGTGAATAGATAAGTCAAAAACCTGTCTATGAGGGGAATGATCACTTTTACCGAGGA 306
|||

QY 241 AAGGCCAGCACTGACACCAATGGGCGGCGCTGCGCTGGAACTCTGCCATGTCTCT 300
|||
Db 307 AAGGCCAGCACTGACACCAATGGGCGGCGCTGCGCTGGAACTCTGCCATGTCTCTCT 366
|||
QY 301 CAGCAACCTTACATATGCCCAAGATCTGATGTCTTCAAGCTGGGCTGGGGAAACATAT 360
|||
Db 367 CAGCAACCTTACATATGCCCAAGATCTGATGTCTTCAAGCTGGGCTGGGGAAACATAT 426
|||
QY 361 TACTCAGGAACCCAGACAAACCGAGGCGACCCCTGTGTCTATGTGACAGTGGGCTTAAG 420
|||
Db 427 TACTCAGGAACCCAGACAAACCGAGGCGACCCCTGTGTCTATGTGACAGTGGGCTTAAG 486
|||
QY 421 CCGCTTGTCCAAAGTGTGATGTGATGATGCTGCGCAATGAAAAAAGCCCTCTCTCTCT 480
|||
Db 487 CCGCTTGTCCAAAGTGTGATGTGATGATGCTGCGCAATGAAAAAAGCCCTCTCTCTCT 546
|||
QY 481 CCAGAAATTAATTAATTTCAATGTGCGCAAAAGCTGTGAGCCCGCTTTAATTAATTT 540
|||
Db 547 CCAGAAATTAATTAATTTCAATGTGCGCAAAAGCTGTGAGCCCGCTTTAATTAATTT 606
|||
QY 541 GGGGAGAAATTCACCAACATGAGAACCAAGCCCTGTGTGGGCGCATCTACAGAGGCGAC 600
|||
Db 607 GGGGAGAAATTCACCAACATGAGAACCAAGCCCTGTGTGGGCGCATCTACAGAGGCGAC 666
|||
QY 601 GGGGGGGGCTGTGTCACTTACGTGTGAGAGGCAAGCTCATACGCCCTTGTGTGTGATC 660
|||
Db 667 GGGGGGGGCTGTGTCACTTACGTGTGAGAGGCAAGCTCATACGCCCTTGTGTGTGATC 726
|||
QY 661 AGCGCCACACCTGCTTATGATTTACCAAGAAAGAGAGACTACATGTCTACCTGGGT 720
|||
Db 727 AGCGCCACACCTGCTTATGATTTACCAAGAAAGAGAGACTACATGTCTACCTGGGT 786
|||
QY 721 CGCTCAAGGCTTAACTTCAACACGCAAGGAGAGTGAAGTTGAGTGAAGAAACCTAATC 780
|||
Db 787 CGCTCAAGGCTTAACTTCAACACGCAAGGAGAGTGAAGTTGAGTGAAGAAACCTAATC 846
|||
QY 781 CTACACAAGGACTACAGGCGTGAACGCTTGTCAACCAAGCAATGTGCTTGTGAAG 840
|||
Db 847 CTACACAAGGACTACAGGCGTGAACGCTTGTCAACCAAGCAATGTGCTTGTGAAG 906
|||
QY 841 ATCCGTTCAGAGGAGGCGAGTGTGCGCAGCACTCCCGGACTATACAGACATCTGCTG 900
|||
Db 907 ATCCGTTCAGAGGAGGCGAGTGTGCGCAGCACTCCCGGACTATACAGACATCTGCTG 966
|||
QY 901 CCTCCGATGATTAACGATTCCTCCAGTGTGCGCAAGCTGTGAGTCACTGCTTTGAAAA 960
|||
Db 967 CCTCCGATGATTAACGATTCCTCCAGTGTGCGCAAGCTGTGAGTCACTGCTTTGAAAA 1026
|||
QY 961 GAGAAATCTACCGACTATCTATCCGAGCAGCTGAAAAATGACTGTGAAAGCTGATT 1020
|||
Db 1027 GAGAAATCTACCGACTATCTATCCGAGCAGCTGAAAAATGACTGTGAAAGCTGATT 1086
|||
QY 1081 TGTGTCTGTGACCCCAAGTGAAGAAACAGATTCTTCCCAAGGAGACTAGAGGGAGCCCTC 1140
|||
Db 1147 TGTGTCTGTGACCCCAAGTGAAGAAACAGATTCTTCCCAAGGAGACTAGAGGGAGCCCTC 1206
|||
QY 1141 GTCTGTTCCCTCCAGGCGGATGACTTTGAGTGTGAGCTGGGGCCGTGATGT 1200
|||
Db 1207 GTCTGTTCCCTCCAGGCGGATGACTTTGAGTGTGAGCTGGGGCCGTGATGT 1266
|||
QY 1201 GCCCTGAAGACAAAGCCAGGCGTCTACACAGAGTCTACACTTTTACCTGGATCCGC 1260
|||
Db 1267 GCCCTGAAGACAAAGCCAGGCGTCTACACAGAGTCTACACTTTTACCTGGATCCGC 1326
|||
QY 1261 AGTACACCAAGAAAGAAATGGCTGGCCCTCTGA 1296
|||
Db 1327 AGTACACCAAGAAAGAAATGGCTGGCCCTCTGA 1362
|||

```
RESULT 5
US-09-466-396A-123
; Sequence 123, Application US/09466396A
; Publication No. US20030119763A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-123

Query Match      100.0%; Score 1296; DB 10; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGCGGCGCTTCTCTGCTGCTGCTGCTGAGCACTCCAAAGC 60
DB 67 ATGAGAGCCCTGCTGCGCGGCGCTTCTCTGCTGCTGCTGCTGAGCACTCCAAAGC 126
QY 61 AGCAATGAATCTTCATCAAGTTCATGAACTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 127 AGCAATGAATCTTCATCAAGTTCATGAACTGCTGCTGCTGCTGCTGCTGCTGCTG 186
QY 121 TCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 187 TCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
QY 181 CACTGTGAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
DB 247 CACTGTGAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 306
QY 241 AAGGCCAGACTGACACCATGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 307 AAGGCCAGACTGACACCATGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
QY 301 CAGCAAACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 360
DB 367 CAGCAAACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 426
QY 361 TACTCAGGAACCCAGACACACCGAGAGGCGACCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 427 TACTCAGGAACCCAGACACACCGAGAGGCGACCTGCTGCTGCTGCTGCTGCTGCTGCT 486
QY 421 CCGCTTGTTCAGAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 487 CCGCTTGTTCAGAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
QY 481 CCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
DB 547 CCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 606
QY 541 GGGGGAGAAATTCACCAACCATGAGAAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 607 GGGGGAGAAATTCACCAACCATGAGAAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 666
QY 601 GGGGGAGAGTCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 667 GGGGGAGAGTCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
QY 661 AGCGCCACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 727 AGCGCCACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
QY 721 CGCTCAAGGCTTAATTCACCAACCGAGAGGAGATGAACTTGAAGTGAAGAAACCTAATC 780
```

```
DB 787 CGCTCAAGGCTTAATTCACCAACCGAGAGGAGATGAACTTGAAGTGAAGAAACCTAATC 846
QY 781 CTACACAAAGACTACAGCGCTGACAGCTTGTGCTACACCAAGCACTTGTGCTGCTGAG 840
DB 847 CTACACAAAGACTACAGCGCTGACAGCTTGTGCTACACCAAGCACTTGTGCTGCTGAG 906
QY 841 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCCGGACTATACAGACCATCTGCTG 900
DB 907 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCCGGACTATACAGACCATCTGCTG 966
QY 901 CCCTGATGTATTAACGATCCCACTTGTGACCAAGCTGATGATCACTGCTTGTGAAAA 960
DB 967 CCCTGATGTATTAACGATCCCACTTGTGACCAAGCTGATGATCACTGCTTGTGAAAA 1026
QY 961 GAGAAATTCACGACTATCTTATCCGAGCAGCTGAAATGACTGTTGTGAAGTGAAT 1020
DB 1027 GAGAAATTCACGACTATCTTATCCGAGCAGCTGAAATGACTGTTGTGAAGTGAAT 1086
QY 1021 TCCCAACGGGAGTGTGAGAGGCGCCACTACTACGGCTGGAAGTCAACCAAAATGCTG 1080
DB 1087 TCCCAACGGGAGTGTGAGAGGCGCCACTACTACGGCTGGAAGTCAACCAAAATGCTG 1146
QY 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGCTGCAAGGAGACTCAGGAGGACCCCTC 1140
DB 1147 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGCTGCAAGGAGACTCAGGAGGACCCCTC 1206
QY 1141 GTCTGTTCCCTCCAAAGCCGCACTGACTTGAATGTGAGCTGCGGCGCTGAGATG 1200
DB 1207 GTCTGTTCCCTCCAAAGCCGCACTGACTTGAATGTGAGCTGCGGCGCTGAGATG 1266
QY 1201 GGCCTGAAGAGCAAGCAGGCGCTTACAGAGAGTCACTCTTACCTGAGATCCCG 1260
DB 1267 GGCCTGAAGAGCAAGCAGGCGCTTACAGAGAGTCACTCTTACCTGAGATCCCG 1326
QY 1261 AGTCAACCCAAAGAGAGATGCGCTGCGCTCTGA 1296
DB 1327 AGTCAACCCAAAGAGAGATGCGCTGCGCTCTGA 1362

RESULT 6
US-10-117-982-123
; Sequence 123, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; NUMBER OF SEQ ID NOS: 2002-04-05
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-123

Query Match      100.0%; Score 1296; DB 14; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGCGGCGCTTCTCTGCTGCTGCTGCTGAGCACTCCAAAGC 60
```


Db	67	ATGAGAGCCTGCTGGGCGGCTGGCTTCTGCGCTCTGGTCGAGACGACTCCAAAGGC	126
OY	61	AGCAATGAACTTCATCAAGTTCATCGAATCTGACTTAAATGAGAGAAATGTGT	120
Db	127	AGCAATGAACTTCATCAAGTTCATCGAATCTGACTTAAATGAGAGAAATGTGT	186
OY	121	TCCAAAGAATCTCTCCAACTTCACTGCTGTGCACTGCCAAAGAAATTCGAGGGCAG	180
Db	187	TCCAAAGAATCTCTCCAACTTCACTGCTGTGCACTGCCAAAGAAATTCGAGGGCAG	246
OY	181	CAGTGAATTAAGTAAATGCAAAACCTGCTATGAGGGGAATGGTCACTTTACCGAGA	240
Db	247	CAGTGAATTAAGTAAATGCAAAACCTGCTATGAGGGGAATGGTCACTTTACCGAGA	306
OY	241	AAGGCCAGCATGACATCAATGGGCGGGCCTGCTGCTGGAATCTGCACTGTCTT	300
Db	307	AAGGCCAGCATGACATCAATGGGCGGGCCTGCTGCTGGAATCTGCACTGTCTT	366
OY	301	CAGCAACGTACCATGCCCAACATCTGATGCTCTTCAGCTGGGCGCTGGGGAACAATAT	360
Db	367	CAGCAACGTACCATGCCCAACATCTGATGCTCTTCAGCTGGGCGCTGGGGAACAATAT	426
OY	361	TACTGACGAACTCCAGACAAACCGAGGCGACCTGTGCTATGTGACAGTGGGCTTAAAG	420
Db	427	TACTGACGAACTCCAGACAAACCGAGGCGACCTGTGCTATGTGACAGTGGGCTTAAAG	486
OY	421	CGCTTGTCGAAGTGCATGTGTGCAATGACTGGGCGATGGAATAAAGCCTCTCTCT	480
Db	487	CGCTTGTCGAAGTGCATGTGTGCAATGACTGGGCGATGGAATAAAGCCTCTCTCT	546
OY	481	CCAGAAATTAATAATTTCAGTGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTAAT	540
Db	547	CCAGAAATTAATAATTTCAGTGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTAAT	606
OY	541	GGGGGGAATTCACCAACATCGAAGAACCGCCCTGTTTGGGGCCATCTACAGAGGCAC	600
Db	607	GGGGGGAATTCACCAACATCGAAGAACCGCCCTGTTTGGGGCCATCTACAGAGGCAC	666
OY	601	CGGGGGGGCTGTGTCACTACGTGTGTGAGGAGAGCCTCATGAGCCTGTGCTGGTGATC	660
Db	667	CGGGGGGGCTGTGTCACTACGTGTGTGAGGAGAGCCTCATGAGCCTGTGCTGGTGATC	726
OY	661	AAGGCGACACACTGCTTCATTGATTAACCCAAAGAGAGAGATCACTGCTTACTGGGT	720
Db	727	AAGGCGACACACTGCTTCATTGATTAACCCAAAGAGAGAGATCACTGCTTACTGGGT	786
OY	721	CGCTCAAGGCTTAATCTCAACAACGCAAGGGAGATTAATTGAGGTGAGAAACCTAATC	780
Db	787	CGCTCAAGGCTTAATCTCAACAACGCAAGGGAGATTAATTGAGGTGAGAAACCTAATC	846
OY	781	CTACACAAAGAGATCAAGCGCTGTACACGCTTGTCTACACACAGACATTTGCTGTGAAG	840
Db	847	CTACACAAAGAGATCAAGCGCTGTACACGCTTGTCTACACACAGACATTTGCTGTGAAG	906
OY	841	ATCCGTTCCAAAGAGGGGAGGTGTGGGACGATCCCGGACATACAGACCATCTGCGCTG	900
Db	907	ATCCGTTCCAAAGAGGGGAGGTGTGGGACGATCCCGGACATACAGACCATCTGCGCTG	966
OY	901	CCCTGATGTATTAAGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGAAAA	960
Db	967	CCCTGATGTATTAAGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGAAAA	1026
OY	961	GAGAAATTCACGACATCTCATCTCGAGAGAGCTGAAATGACATGTTGTGAAGCTGAT	1020
Db	1027	GAGAAATTCACGACATCTCATCTCGAGAGAGCTGAAATGACATGTTGTGAAGCTGAT	1086
OY	1021	TCCACCGGGAGTGTACAGAGCCCACTACTACGGCTCTGAAGTACCAACCAAAATGCTG	1080
Db	1087	TCCACCGGGAGTGTACAGAGCCCACTACTACGGCTCTGAAGTACCAACCAAAATGCTG	1146
OY	1081	TGTGTGTGACCCCAAGTGGAAAAAGATTTCTGTCAGGGAGACTCAGGGGACCTCTT	1140

Db	1147	TGTCGTCTGACCCACAGTGGAAAAAGATTCTCCAGGGAGACTCGAGGGGAGACCCCTC	1206
Qy	1141	GTCTGTTCCCTCCAAAGGCGCGATGACTTTGACTGGAAATTGTAGCTGGAGCGGTGGATGT	1200
Db	1207	GTCTGTTCCCTCCAAAGGCGCGATGACTTTGACTGGAAATTGTAGCTGGAGCGGTGGATGT	1266
Qy	1201	GCCCTGAGAGCAAGCCAGCGCTTACACGAGAGTCTACACTTTTACCTTGAATCCG	1260
Db	1267	GCCCTGAGAGCAAGCCAGCGCTTACACGAGAGTCTACACTTTTACCTTGAATCCG	1328
Qy	1261	AGTCACACCAAGAAAGAAATGGCGCCCTCTGA	1296
Db	1327	AGTCACACCAAGAAAGAAATGGCGCCCTCTGA	1362
RESULT 7			
US-10-313-986-123			
/ Sequence 123, Application US/10313986			
/ Publication No. US20030236209A1			
/ GENERAL INFORMATION:			
/ APPLICANT: Foy, Teresa M.			
/ APPLICANT: McNabb, Andria			
/ APPLICANT: Watanabe, Yoshihiro			
/ APPLICANT: Reed, Steven G.			
/ APPLICANT: Wang, Tongrong			
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER			
/ FILE REFERENCE: 210121.455C19			
/ CURRENT APPLICATION NUMBER: US/10/313,986			
/ CURRENT FILING DATE: 2002-12-04			
/ NUMBER OF SEQ ID NOS: 560			
/ SOFTWARE: FastSeq for Windows Version 4.0			
/ SEQ ID NO 123			
/ LENGTH: 2294			
/ TYPE: DNA			
/ ORGANISM: Homo sapiens			
US-10-313-986-123			
Query Match			
Best Local Similarity 100.0%; Score 1296; DB 15; Length 2294;			
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGAGAGCCCTGCTGGCGCGCTCTCTCTGCGCTCTGCTGTGAGCGACTCCAAAGC	60
Db	67	ATGAGAGCCCTGCTGGCGCGCTCTCTCTGCTGTGCTGCTGTCGTGAGCGACTCCAAAGC	126
Qy	61	AGCAATGAACCTTCAATCAAGTTCATGCAACTGTCGACCTGTCTAAATGGAGAACTGTGTG	120
Db	127	AGCAATGAACCTTCAATCAAGTTCATGCAACTGTCGACCTGTCTAAATGGAGAACTGTGTG	186
Qy	121	TCCAACAAGTACTTCTCCAACTTCACTGTGTGCAACTGCCAAGAANAATTCGAGGGCAG	180
Db	187	TCCAACAAGTACTTCTCCAACTTCACTGTGTGCAACTGCCAAGAANAATTCGAGGGCAG	246
Qy	181	CACGTGTAAATAGTAAGTCAAAAACTGCTATGAGGGGAATGTCTCACTTTTACCGAGGA	240
Db	247	CACGTGTAAATAGTAAGTCAAAAACTGCTATGAGGGGAATGTCTCACTTTTACCGAGGA	306
Qy	241	AAGGCCAGCACTGACCAATGAGGCGCGCTCTGCTGCTCCAGAACTGTGCCACTGTCTT	300
Db	307	AAGGCCAGCACTGACCAATGAGGCGCGCTCTGCTGCTCCAGAACTGTGCCACTGTCTT	366
Qy	301	CAGCAAACTGACATGCTCCACAGATCTGATGCTCTTCACTGAGGCTGGGGAAACATAAT	360
Db	367	CAGCAAACTGACATGCTCCACAGATCTGATGCTCTTCACTGAGGCTGGGGAAACATAAT	426
Qy	361	TATCGCAGGAACCCAGACCAACCGGAGGCGACCCGTGGTGCTATGTGAGAGTGGGCTTAAAG	420
Db	427	TATCGCAGGAACCCAGACCAACCGGAGGCGACCCGTGGTGCTATGTGAGAGTGGGCTTAAAG	486
Qy	421	CCGCTTGTCCAAAGTGCATGTGTGACTGTGCGAGATGGAAAAAGCCCTTCTCTCT	480
Db	487	CCGCTTGTCCAAAGTGCATGTGTGACTGTGCGAGATGGAAAAAGCCCTTCTCTCT	546

QY 481 CCAGAGAAATTAATTTAGTGTGGCCAAAGACTGTAGGCCCCGCTTAAGATTAT 540
 DB 547 CCGAGAAATTAATTTAGTGTGGCCAAAGACTGTAGGCCCCGCTTAAGATTAT 606
 QY 541 GGGGGGAAATTCACCAATCGAAGAACAGCCCTGTGTTGGCCATCTACAGAGGAC 600
 DB 607 GGGGGGAAATTCACCAATCGAAGAACAGCCCTGTGTTGGCCATCTACAGAGGAC 666
 QY 601 GGGGGGGGCTGTGTACCTACGTTGTGTGAGGAGGAGCTTACAGCCCTTGTGGTATC 660
 DB 667 GGGGGGGGCTGTGTACCTACGTTGTGTGAGGAGGAGCTTACAGCCCTTGTGGTATC 726
 QY 661 AGGCGCACACACTGTCTTCAATTAATCCCAAGAGAGAGACTACATCGCTTACCTGGGT 720
 DB 727 AGGCGCACACACTGTCTTCAATTAATCCCAAGAGAGAGACTACATCGCTTACCTGGGT 786
 QY 721 CGCTCAAGGCTTAATCTTCAACACGCAAGGGAGATGAAATTTGAGGTGAAAACCTAATC 780
 DB 787 CGCTCAAGGCTTAATCTTCAACACGCAAGGGAGATGAAATTTGAGGTGAAAACCTAATC 846
 QY 781 CTACACAGAGACTACAGCGCTGACAGCTTGTCTACCAACGACATTTGCTTGTGAAG 840
 DB 847 CTACACAGAGACTACAGCGCTGACAGCTTGTCTACCAACGACATTTGCTTGTGAAG 906
 QY 841 ATCCGTTCCAAAGAGGGGAGGTGTGGGAGCCATCCCGGACTATACAGACATCTGCTG 900
 DB 907 ATCCGTTCCAAAGAGGGGAGGTGTGGGAGCCATCCCGGACTATACAGACATCTGCTG 966
 QY 901 CCTCGATGTATACGATCCCGAGTTTGGGACAGCTGTGAGATCACTGGCTTTGGAAAA 960
 DB 967 CCTCGATGTATACGATCCCGAGTTTGGGACAGCTGTGAGATCACTGGCTTTGGAAAA 1026
 QY 961 GAGAAATTTCAACGACTATCTTATCCGAGAGAGCTGAAAATGACTGTGTGAAGCTGATT 1020
 DB 1027 GAGAAATTTCAACGACTATCTTATCCGAGAGAGCTGAAAATGACTGTGTGAAGCTGATT 1086
 QY 1021 TCCGACCGGGAGGTGTAGAGAGAGCCGCACTATACGGGCTGTGAATCAACCAAAATGCTG 1080
 DB 1087 TCCGACCGGGAGGTGTAGAGAGAGCCGCACTATACGGGCTGTGAATCAACCAAAATGCTG 1146
 QY 1081 TGTGTGTGTGACCCACAGTGTGAAGAAAGATTTCTGTGACAGGAGACTCAAGGGGAGCCCTC 1140
 DB 1147 TGTGTGTGTGACCCACAGTGTGAAGAAAGATTTCTGTGACAGGAGACTCAAGGGGAGCCCTC 1206
 QY 1141 GTCTGTTCCTTCCAAAGCCGCGATGACTTTGACTGAATGTGTGAGCTGTGGGCGCTGTGATGT 1200
 DB 1207 GTCTGTTCCTTCCAAAGCCGCGATGACTTTGACTGAATGTGTGAGCTGTGGGCGCTGTGATGT 1266
 QY 1201 GGCCTGAAGAGACAGGAGGAGGCTGTACAGAGAGTCTCAACCTTTTACCTTGTGATCCGC 1260
 DB 1267 GGCCTGAAGAGACAGGAGGAGGCTGTACAGAGAGTCTCAACCTTTTACCTTGTGATCCGC 1326
 QY 1261 AGTCAACACCAAGAGAGAGATGAGCTGTGGCCCTCTGA 1296
 DB 1327 AGTCAACACCAAGAGAGAGATGAGCTGTGGCCCTCTGA 1362

; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Fanger, Neil
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C14
 ; CURRENT APPLICATION NUMBER: US/09/735,705
 ; CURRENT FILING DATE: 2000-12-12
 ; NUMBER OF SEQ ID NOS: 419
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 122
 ; LENGTH: 1475
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-735-705-122
 Query Match 99.8%; Score 1292.8; DB 9; Length 1475;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGAGAGCCCTGCTGGCGGGCCCTGCTTCTGTGGTCTGTGGTGTGAGGAGCTCAAGGC 60
 DB 81 ATGAGAGCCCTGCTGGCGGGCCCTGCTTCTGTGGTCTGTGGTGTGAGGAGCTCAAGGC 140
 QY 61 AGCAATGAATTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCAT 120
 DB 141 AGCAATGAATTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCAT 200
 QY 121 TCCAAAGTACTTCTTCCAAATTCATGAGTGTGCACTGCTCCAAAGAAATTTGGAGGGCAG 180
 DB 201 TCCAAAGTACTTCTTCCAAATTCATGAGTGTGCACTGCTCCAAAGAAATTTGGAGGGCAG 260
 QY 181 CACTGTGAATTAATTAAGTCAAAACCTGTGTATGAGGGGAAATGTCACTTTTACCGAGGA 240
 DB 261 CACTGTGAATTAATTAAGTCAAAACCTGTGTATGAGGGGAAATGTCACTTTTACCGAGGA 320
 QY 241 AAGGCGAGAGCTGACACCAATGAGGCGGGCCCTGCTGCTGCACTGTGCTT 300
 DB 321 AAGGCGAGAGCTGACACCAATGAGGCGGGCCCTGCTGCTGCACTGTGCTT 380
 QY 301 CAGCAAACTGACATGCGCAGATGATGCTTCACTGAGCTGTGGGCTGTGGGAAACATTAAT 360
 DB 381 CAGCAAACTGACATGCGCAGATGATGCTTCACTGAGCTGTGGGCTGTGGGAAACATTAAT 440
 QY 361 TACTGAGAGAACCCAGACCAACCGAGGCGACCTGTGTGTATGTGAGGTGTGGCTTAAAG 420
 DB 441 TACTGAGAGAACCCAGACCAACCGAGGCGACCTGTGTGTATGTGAGGTGTGGCTTAAAG 500
 QY 421 CCGCTGTGCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 DB 501 CCGCTGTGCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 560
 QY 481 CCGAGAGATTAATTAATTTAGTGTGGCCAAAGACTGTGAGGCCCTTTAAGATTAT 540
 DB 561 CCGAGAGATTAATTAATTTAGTGTGGCCAAAGACTGTGAGGCCCTTTAAGATTAT 620
 QY 541 GGGGGGAAATTCACCAATCGAAGAACAGCCCTGTGTTGGCCATCTACAGAGGAC 600
 DB 621 GGGGGGAAATTCACCAATCGAAGAACAGCCCTGTGTTGGCCATCTACAGAGGAC 680
 QY 601 GGGGGGGGCTGTGTACCTACGTTGTGTGAGGAGGAGCTGATGAGCCCTTGTGGTATC 660
 DB 681 GGGGGGGGCTGTGTACCTACGTTGTGTGAGGAGGAGCTGATGAGCCCTTGTGGTATC 740
 QY 661 AGGCGCACACACTGTCTTCAATTAATCCCAAGAGAGAGACTACATCGCTTACCTGGGT 720
 DB 741 AGGCGCACACACTGTCTTCAATTAATCCCAAGAGAGAGACTACATCGCTTACCTGGGT 800
 QY 721 CGCTCAAGGCTTAATCTTCAACACGCAAGGGAGATGAAATTTGAGGTGAAAACCTAATC 780
 DB 801 CGCTCAAGGCTTAATCTTCAACACGCAAGGGAGATGAAATTTGAGGTGAAAACCTAATC 860
 QY 781 CTACACAGAGACTACAGCGCTGACAGCTTGTCTACCAACGACATTTGCTTGTGAAG 840

RESULT 8

US-09-735-705-122.
 ; Sequence 122, Application US/09735705
 ; Patent No. US2002005239A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hoeken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aljun
 ; APPLICANT: Skeiky, Yaelir A.W.
 ; APPLICANT: Henderson, Robert A.

Db	861	CTACAGACAGACTACAGCCCTGACACGCTTTGCTCACACACACAGCAATTGCTTGTAAG	920
Qy	841	ATCCGTTCCAGAGGGGAGGTGTGTGGACGCCATCCCGACTATACAGACATCTGCTG	900
Db	921	ATCCGTTCCAGAGGGGAGGTGTGTGGACGCCATCCCGACTATACAGACATCTGCTG	980
Qy	901	CCCTGATGTATTAACGATCCCAAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGAAAA	960
Db	981	CCCTGATGTATTAACGATCCCAAGTTTGGCACAAGCTGTGAGATCACTGGCTTTGAAAA	1040
Qy	961	GAGAAATTCACCGACTATCTATCCGAGCAGCTGAAATAAGTACTGTTGTGAAGCTGATT	1020
Db	1041	GAGAAATTCACCGACTATCTATCCGAGCAGCTGAGATATCTGTTGTGAAGCTGATT	1100
Qy	1021	TCCCAACCGGAGTGTCCAGACGCCCACTATACGACTGTGAAGTACCAACAAATGCTG	1080
Db	1101	TCCCAACCGGAGTGTCCAGACGCCCACTACTACGGCTGTGAAGTACCAACAAATGCTG	1160
Qy	1081	TGTGCTGCTGACCCACAGTGGAAAAACAATTCTCTCCAGGGAGATTCAGGGGGACCCCTC	1140
Db	1161	TGTGCTGCTGACCCACAGTGGAAAAACAATTCTCTCCAGGGAGATTCAGGGGGACCCCTC	1220
Qy	1141	GTCTGTTCCTTCMAAGCCCGCATGACTTTGACTGGAAATGTGAGCTGAGGCCGTGAGATGT	1200
Db	1221	GTCTGTTCCTTCMAAGCCCGCATGACTTTGACTGGAAATGTGAGCTGAGGCCGTGAGATGT	1280
Qy	1201	GCCCTGGAGACCAACCCAGCGGCTCTACACGAGAGTCTACACTTTCTTACCCTGGATCCG	1260
Db	1281	GCCCTGGAGAGCAACCCAGCGGCTCTACACGAGAGTCTACACTTTCTTACCCTGGATCCG	1340
Qy	1261	AGTCACACCAAGAGAGATAGCTGGGCCCTCTGA	1296
Db	1341	AGTCACACCAAGAGAGATAGCTGGGCCCTCTGA	1376
RESULT 9			
US-09-850-716A-122			
; Sequence 122, Application US/09850716A			
; Patent No. US20020115139A1			
; GENERAL INFORMATION:			
; APPLICANT: Kaloe, Michael D.			
; APPLICANT: McNeill, Patricia D.			
; APPLICANT: Retter, Marc W.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; FILE REFERENCE: 210121.455C15			
; CURRENT APPLICATION NUMBER: US/09/850.716A			
; NUMBER OF SEQ ID NOS: 440			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 122			
; LENGTH: 1475			
; TYPE: DNA			
; ORGANISM: Homo sapien			
US-09-850-716A-122			

Query Match	99.8%	Score 1292.8	DB 9	Length 1475
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1294	Conservative	0	Mismatches 2	Indels 0
			Gaps 0	
QY	1 ATGAGAGCCCTGCTGGCGCGCCTGTTCTCTGCGCGTCTGGTGTGAGGAGCATCCCAAAAGC	60		
Db	81 ATGAGAGCCCTGCTGGCGCGCCTGTTCTCTGCGCGTCTGGTGTGAGGAGCATCCCAAAAGC	140		
QY	61 AGCATGAATTTTCATCAAGTTTCATCGAATCTGTGACTCTTAATATGAGAGAACTGTGTG	120		
Db	141 AGCATGAATTTTCATCAAGTTTCATCGAATCTGTGACTCTTAATATGAGAGAACTGTGTG	200		
QY	121 TCCAAACAAGTACTTCTCCAACTTTCATCTGCTGCAACCTCCCAAGAAATTGGAGGGGAG	180		
Db	201 TCCAAACAAGTACTTCTCCAACTTTCATCTGCTGCTGCAACCTCCCAAGAAATTGGAGGGGAG	260		
QY	181 CACTGTGAATATGATTAAGTCAAAAAACCTGCTATGAGGGGAAATGTCACTTTTACCGAGGA	240		

Db	261	CACCTGGAATTAGTAAAGTCAAAAAACCTGCTATGAGAGGAAATGGTCACTTTTAA	CCAGAGA	320
Qy	241	AAGGCCAGCACTGACACATGggGCCGCCCTGCTGCTCCCTGGAACTCTGCACATGCTT		300
Db	321	AAGGCCAGCACTGACACATGggGCCGCCCTGCTGCTCCCTGGAACTCTGCACATGCTT		380
Qy	301	CAGCAAAAGCTAACATGCCCAAGATCTGATGCTCTTCACTGggGCTGGGAAACAATAT		360
Db	381	CAGCAAAAGCTAACATGCCCAAGATCTGATGCTCTTCACTGggGCTGGGAAACAATAT		440
Qy	361	TACTGAGGAACCCAGACAACCGAGGCGAACCCCTGGTGTATATGTGAGGTGGGCTTAAG		420
Db	441	TACTGAGGAACCCAGACAACCGAGGCGAACCCCTGGTGTATATGTGAGGTGGGCTTAAG		500
Qy	421	CCGCTTGTCCAAAGATGTGACATGgTGCATGACTGCGCAGATGGA	AAAAAGCCCTCTCTCT	480
Db	501	CCGCTTGTCCAAAGATGTGACATGgTGCATGACTGCGCAGATGGA	AAAAAGCCCTCTCTCT	560
Qy	481	CCAGAAAGATTTAAAAATTTCAGTGTGGCCAAAAGACCTGAGGGCCCGGCTTTAAATTA	TTT	540
Db	561	CCAGAAAGATTTAAAAATTTCAGTGTGGCCAAAAGACCTGAGGGCCCGGCTTTAAATTA	TTT	620
Qy	541	GGGGAGAAATTCACCAACATCGAGAAACAAGCCCTGGTGTTCGGCCATCTACAGAGGAC		600
Db	621	GGGGAGAAATTCACCAACATCGAGAAACAAGCCCTGGTGTTCGGCCATCTACAGAGGAC		680
Qy	601	GGGGGGGGGCTGTGCAACCTGAGCTGTGTGAGGCAACCTCATCAGCCCTTGGGCTGATC		660
Db	681	GGGGGGGGGCTGTGCAACCTGAGCTGTGTGAGGCAACCTCATCAGCCCTTGGGCTGATC		740
Qy	661	AGCGCACAACACTGTCTTATGATTATCCAAAGAAAGAGAGACTACATGCTCACTGGGT		720
Db	741	AGCGCACAACACTGTCTTATGATTATCCAAAGAAAGAGAGACTACATGCTCACTGGGT		800
Qy	721	CGCTCAAGGCTTAATCTCCAAACCGAAGGGAGATGAAGTTGAGGTGAAAACCTAATC		780
Db	801	CGCTCAAGGCTTAATCTCCAAACCGAAGGGAGATGAAGTTGAGGTGAAAACCTAATC		860
Qy	781	CTACACAAGACTACAGGCTGACACGGTGTCTACCAACAAGCATGTCCTGCTGAG		840
Db	861	CTACACAAGACTACAGGCTGACACGGTGTCTACCAACAAGCATGTCCTGCTGAG		920
Qy	841	ATCGCTTCCAAAGAGGAGGAGGTGTGCGAGCCATCCCGACTATACAGCCATCTGCTG		900
Db	921	ATCGCTTCCAAAGAGGAGGAGGTGTGCGAGCCATCCCGACTATACAGCCATCTGCTG		980
Qy	901	CCCTGATGTATAACGATCCCAAGTTTGGCAACAGCTGTGAGTACATGGCTTTGAAAA		960
Db	981	CCCTGATGTATAACGATCCCAAGTTTGGCAACAGCTGTGAGTACATGGCTTTGAAAA		1040
Qy	961	GAGAAATTCTACACTATCTCTATCCGAGCAGCTGAAATAGACTGTGTGAACTGAT		1020
Db	1041	GAGAAATTCTACACTATCTCTATCCGAGCAGCTGAAATAGACTGTGTGAACTGAT		1100
Qy	1021	TCCCAACCGGAGGTGTGACAGCCCCCACTACAGGCTCTGAAGTCAACA	CCAAATGCTG	1080
Db	1101	TCCCAACCGGAGGTGTGACAGCCCCCACTACAGGCTCTGAAGTCAACA	CCAAATGCTG	1160
Qy	1081	TGTGCTGTGTACCCACAGTGTGAAAACAGATTCTGCGACAGGAGACTCAGGGGGACCCCTC		1140
Db	1161	TGTGCTGTGTACCCACAGTGTGAAAACAGATTCTGCGACAGGAGACTCAGGGGGACCCCTC		1220
Qy	1141	GTCTGTTCCTTCCAAAGGCCGATACCTTTGACTGGAATGTGTGAGCTGAGGCGCTGTAGT		1200
Db	1221	GTCTGTTCCTTCCAAAGGCCGATACCTTTGACTGGAATGTGTGAGCTGAGGCGCTGTAGT		1280
Qy	1201	GGCGTGAAGGACAAGCCAGGCGTCTTACAAGAGTCTCAACTTTTAACTCTGATTCGC		1260
Db	1281	GGCGTGAAGGACAAGCCAGGCGTCTTACAAGAGTCTCAACTTTTAACTCTGATTCGC		1340
Qy	1261	AGTCACACAAGAGAAATGCGCTGCGCCCTGTGA		1296

[illegible]

```

RESULT 13
US-10-117-982-122
; Sequence 122: Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darlick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117.982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-122

Query Match      99.8%; Score 1292.8; DB 14; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1      ATGAGAGCCCTGCTGGCGCGCCCTGCTCTCTGCGGTCGATCGTAGGCACTCCAAAGC 60
DB      81      ATGAGAGCCCTGCTGGCGCGCCCTGCTCTCTGCGGTCGATCGTAGGCACTCCAAAGC 140

QY      61      AGCAATGAACCTTCATCAAGTTCATCGAATGTGACTGTCTAAATGGAGAAATGTTGTG 120
DB      141     AGCAATGAACCTTCATCAAGTTCATCGAATGTGACTGTCTAAATGGAGAAATGTTGTG 200

QY      121     TTCACAAGTAATCTTCCAAATTCATCTGTGTGCAATGCCCAAGAAATTTGGAGGGCAG 180
DB      201     TTCACAAGTAATCTTCCAAATTCATCTGTGTGCAATGCCCAAGAAATTTGGAGGGCAG 260

QY      181     CACTGTAAATAGTAAGTCAAAAACCTGTGATAGAGGGAAATGTGCACTTTTACCGAGA 240
DB      261     CACTGTAAATAGTAAGTCAAAAACCTGTGATAGAGGGAAATGTGCACTTTTACCGAGA 320

QY      241     AAGCCAGCACTGACACCATAGGCGCGCCCTGCTGCTCCCTGGAACCTGTGCATGTCTT 300
DB      321     AAGCCAGCACTGACACCATAGGCGCGCCCTGCTGCTCCCTGGAACCTGTGCATGTCTT 380

QY      301     CAGGAAACGTATCAATGCGCCAGATCTGATGCTCTTACGCTGGGCGCTGGGGAAACATAT 360
DB      381     CAGGAAACGTATCAATGCGCCAGATCTGATGCTCTTACGCTGGGCGCTGGGGAAACATAT 440

QY      361     TACTGCAAGAACCCAGACAACCCGAGGCGCACTGTGTCTATGTGCAAGTGGGCTTAAG 420
DB      441     TACTGCAAGAACCCAGACAACCCGAGGCGCACTGTGTCTATGTGCAAGTGGGCTTAAG 500

QY      421     CCGCTTGTCCAGAGTGTGATGTGATGATCTGCGCAATGGAATAAAGCCCTCTCTCTCT 480
DB      501     CCGCTTGTCCAGAGTGTGATGTGATGATCTGCGCAATGGAATAAAGCCCTCTCTCTCT 560

QY      481     CCAGAAAGATTTAAATTTCAAGTGTGGCCAAAGACTCTGAGGGCCCGCTTTAAATATT 540
DB      561     CCAGAAAGATTTAAATTTCAAGTGTGGCCAAAGACTCTGAGGGCCCGCTTTAAATATT 620

QY      541     GGGGAGAAATTCACCAACATGAGAACGAGCCCTGTGTTTGGCGCCATCTACAGAGGAC 600

```


Dp	62	GGGGGAGAAATTCAACCAACATCCAGAAACAGCCCTGGTTTGGCGGCATTTACAGAGGCAC	660
Qy	601	CGGGGGGGCTCTGTCACTACGTGTGTGAGGCGAGCTTCATAGCCCTTGTCTGGTGATC	660
Dp	681	CGGGGGGGGCTCTGTCACTACGTGTGTGAGGCAAGCTTCATAGCCCTTGTCTGGTGATC	740
Qy	661	AGCCCAACACAGTCTTCAATTGATTTCCMAAGAAAGAGAGAGTCACTATCTACCTGGGT	720
Dp	741	AGGCGCAACACAGTCTTCAATTGATTTCCMAAGAAAGAGAGAGTCACTATCTACCTGGGT	800
Qy	721	CGCTCAAGGCTTAACTTCCAAACACGCAAGGGAGATGAAGTTTGAAGTGGAAACCTATC	780
Dp	801	CGCTCAAGGCTTAACTTCCAAACGCAAGGGAGATGAAGTTTGAAGTGGAAACCTATC	860
Qy	781	CTACACAAGAGCTACAGCGCTGACAGCTTGTCAACCAACGACATTTGCTGTGAAG	840
Dp	861	CTACACAAGAGCTACAGCGCTGACAGCTTGTCAACCAACGACATTTGCTGTGAAG	920
Qy	841	ATCCGTTTCAAGAGGGGAGAGTGTGCGAGCCATCCCGACTATACAGACCATCTGCCTG	900
Dp	921	ATCCGTTTCAAGAGGGGAGAGTGTGCGAGCCATCCCGACTATACAGACCATCTGCCTG	980
Qy	901	CCCCCAATGTATACGATCCCAAGTTTGGCACAAAGCTGTGAGATCACTGCCTTTGAAAA	960
Dp	981	CCCCCAATGTATACGATCCCAAGTTTGGCACAAAGCTGTGAGATCACTGCCTTTGAAAA	1040
Qy	961	GAGAAATTCTAACCGACTATCTATCCGAGAGCAAGCTGAAAAATGACTGTTTGAAGCTGATT	1020
Dp	1041	GAGAAATTCTAACCGACTATCTATCCGAGAGCAAGCTGAAAAATGACTGTTTGAAGCTGATT	1100
Qy	1021	TCCACACGGGAGGTGTACAGAGGCCCAACTACACGGCTTGAAGTCAACCAAAATGCTG	1080
Dp	1101	TCCACACGGGAGGTGTACAGAGGCCCAACTACACGGCTTGAAGTCAACCAAAATGCTG	1160
Qy	1081	TGTGCTGTGACCAACAGTGGAAAAACAGATTCTGTCCAGGAGAGACTCAGGGGGAGCCCTC	1140
Dp	1161	TGTGCTGTGACCAACAGTGGAAAAACAGATTCTGTCCAGGAGAGACTCAGGGGGAGCCCTC	1220
Qy	1141	GTCGTGTCCCTCCACAGGCCGCACTGACTTTGACTGGAATTGTGAGCTGGGGCGTGGATGT	1200
Dp	1221	GTCGTGTCCCTCCACAGGCCGCACTGACTTTGACTGGAATTGTGAGCTGGGGCGTGGATGT	1280
Qy	1201	GCCCTGAAGAGCAAGCGCAGGCTGTACACGAGAGTCTCACACTTCTTAAACCTGGATCCGC	1260
Dp	1281	GCCCTGAAGAGCAAGCGCAGGCGTGTACACGAGAGTCTCACACTTCTTAAACCTGGATCCGC	1340
Qy	1261	AGTCAACACCAAGAAAGAAATGCGCTTGGCCTCTTGA	1296
Dp	1341	AGTCAACACCAAGAAAGAAATGCGCTTGGCCTCTTGA	1376

```

RESULT 14
US-10-101-510-159
; Sequence 159, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: MAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 159
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-159

Query Match          99.8%; Score 1292.8; DB 14; Length 1475

```

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY	61	ATGAGAGCCCTGCTGGGCGGGCCGCTTCTCTGGGCTCTGGTCTGTGAGCCACTCCAAAGC	60
Db	81	ATAGAGGCCCTGCTGGGCGGGCCGCTTCTCTGGGCTCTGGTCTGTGAGCCACTCCAAAGC	140
QY	61	AGCAATGAACTTCATCAAGTTCCATCGAACTGTGACTGTCTAAATGAGAGAACATGTGT	120
Db	141	AGCAATGAACTTCATCAAGTTCCATCGAACTGTGACTGTCTAAATGAGAGAACATGTGT	200
QY	121	TCGAACAATTAATCTTCCCAACAATTGACTGTGTGAACTGCCCAAGAAATTCCGAGGCGAG	180
Db	201	TCGAACAATTAATCTTCCCAACAATTGACTGTGTGAACTGCCCAAGAAATTCCGAGGCGAG	260
QY	181	CACTGTGAAATGATTAAGTCAAAAACTGTCTATGAGGGGAATAGTCACTTTTACCGAGA	240
Db	261	CACTGTGAAATGATTAAGTCAAAAACTGTCTATGAGGGGAATAGTCACTTTTACCGAGA	320
QY	241	AAGGCCAGACTGACCACTAGGGCGGGCCCTGCTGCTGCCGAACTGTGCTT	300
Db	321	AAGGCCAGACTGACCACTAGGGCGGGCCCTGCTGCTGCCGAACTGTGCTT	380
QY	301	CAGCAAACTATCATGCTCCACAGATCTGATGTCTTCACTGAGCTGGGGCTGGGAAACAAT	360
Db	381	CAGCAAACTATCATGCTCCACAGATCTGATGTCTTCACTGAGCTGGGGCTGGGAAACAAT	440
QY	361	TACTGAGGAACCCAGACACCGGAGGGGACCTCTGTGCTATGTGCAAGGTGGGCTTAAAG	420
Db	441	TACTGAGGAACCCAGACACCGGAGGGGACCTCTGTGCTATGTGCAAGGTGGGCTTAAAG	500
QY	421	CCGCTGTGTCAAGAGTGCATGTGTGATGACTGCGAGATGAGAAAAAGCCCTCTCTCCT	480
Db	501	CCGCTGTGTCAAGAGTGCATGTGTGATGACTGCGAGATGAGAAAAAGCCCTCTCTCCT	560
QY	481	CCAGAAGAAATTAATTTCAAGTGTGGCCAAAGAACTGTGAGGCGCCGCTTTAAGATTAAT	540
Db	561	CCAGAAGAAATTAATTTCAAGTGTGGCCAAAGAACTGTGAGGCGCCGCTTTAAGATTAAT	620
QY	541	GGGGGAGAAATTCACCAAGCATGAGAACAGAGCCCTGTGTGGGGCATCTACAGAGGAGAC	600
Db	621	GGGGGAGAAATTCACCAAGCATGAGAACAGAGCCCTGTGTGGGGCATCTACAGAGGAGAC	680
QY	601	CGGGGGGGCTCTGTCACTAGCTGTGTGTGAGAGGACCTCATCAGCCCTTGTGGGTGATC	660
Db	681	CGGGGGGGCTCTGTCACTAGCTGTGTGTGAGAGGACCTCATCAGCCCTTGTGGGTGATC	740
QY	661	AGCGGCACACCTGCTTCAATGATTAATCCCAAGAAAGAGACTACATGTCTACTCTGGGT	720
Db	741	AGCGGCACACCTGCTTCAATGATTAATCCCAAGAAAGAGACTACATGTCTACTCTGGGT	800
QY	721	CGCTCAAGGCTTAACTCCACACGCAAGGAGGAGATGAAGTTTGAAGTGAAGAAACCTAATC	780
Db	801	CGCTCAAGGCTTAACTCCACACGCAAGGAGGAGATGAAGTTTGAAGTGAAGAAACCTAATC	860
QY	781	CTACACAAGAGCTACAGGCTGACAGAGCTTGTCTCAACCAAGACATTTGCTTGTCTGAAG	840
Db	861	CTACACAAGAGCTACAGGCTGACAGAGCTTGTCTCAACCAAGACATTTGCTTGTCTGAAG	920
QY	841	ATCCGTTTCAAGAGGAGGAGGTGTGCGCAGCATCCCGGACTTATACAGACCATCTGCGCTG	900
Db	921	ATCCGTTTCAAGAGGAGGAGGTGTGCGCAGCATCCCGGACTTATACAGACCATCTGCGCTG	980
QY	901	CCCTCGATGTATAAGATCCCAAGTTTGGCAAGACTGTGAGATCACTGCTTTGAGAAA	960
Db	981	CCCTCGATGTATAAGATCCCAAGTTTGGCAAGACTGTGAGATCACTGCTTTGAGAAA	1040
QY	961	GAGAAATTCACCGACTATCTATCCCGGAGAGCTGAAAAATGATGTGTGAGAGCTAAT	1020
Db	1041	GAGAAATTCACCGACTATCTATCCCGGAGAGCTGAAAAATGATGTGTGAGAGCTAAT	1100
QY	1021	TCCCAACGGGAGTGTCAAGACGCCCACTACTACGGCTTGAAGTCAACACCAAAATGCTG	1080


```

Db      1101 TCCACCGGAGTGTGACGAGCCCACTACTAGGCTGTGAGTCAACCAACAAATCTGT 1160
Qy      1081 TGTGCTGTGACCCACAGTGTGAAAACAGATTTCTGTCCAGGAGACTCAAGGGGAAACCCCTC 1140
Db      1161 TGTGCTGTGACCCACAGTGTGAAAACAGATTTCTGTCCAGGAGACTCAAGGGGAAACCCCTC 1220
Qy      1141 GTCTGTTCCTCCAGGCGGCATGACTTTGACTGAAATTTGAGTGTGGGCGCGTGAATGT 1200
Db      1221 GTCTGTTCCTCCAGGCGGCATGACTTTGACTGAAATTTGAGTGTGGGCGCGTGAATGT 1280
Qy      1201 GCCCTGAAGAGACAGCCAGGCTGTACAGAGAGTCCACACTTTTACCTTGATCCGC 1260
Db      1281 GCCCTGAAGAGACAGCCAGGCTGTACAGAGAGTCCACACTTTTACCTTGATCCGC 1340
Qy      1261 AGTCACACCAAGAAAGAAATGGCTGTGGCCCTCTGA 1296
Db      1341 AGTCACACCAAGAAAGAAATGGCTGTGGCCCTCTGA 1376

RESULT 15
US-10-313-986-122
; Sequence 122, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-122

Query Match      99.8%; Score 1292.8; DB 15; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATGAGAGCCCTGTGCGCGCCCTGTCTCTGCGTCCGTGCGTGAAGCACTCAAGGC 60
Db      81 ATGAGAGCCCTGTGCGCGCCCTGTCTCTGCGTCCGTGCGTGAAGCACTCAAGGC 140
Qy      61 AGCAATGAACCTTCATCAAGTTCATGAACTGTGACTGTCTAAATGAGAGAAATGTGTG 120
Db      141 ACCAAATGAACCTTCATCAAGTTCATGAACTGTGACTGTCTAAATGAGAGAAATGTGTG 200
Qy      121 TCCAAAGATTAATCTTCCCAACATTTCACTGTGCAACTGCCCCAAAGAAATTCGAGAGGCGAG 180
Db      201 TCCAAAGATTAATCTTCCCAACATTTCACTGTGCAACTGCCCCAAAGAAATTCGAGAGGCGAG 260
Qy      181 CACTGTGAATAGATAAGTCAAAAACCTGCTATGAGGGGAAATGTAACCTTTTACCGAGA 240
Db      261 CACTGTGAATAGATAAGTCAAAAACCTGCTATGAGGGGAAATGTAACCTTTTACCGAGA 320
Qy      241 AAGGCCAGCACTGACACCAATGAGGCGCGCCCTGTGCTGCAACTCTGTGCACTGTCTT 300
Db      321 AAGGCCAGCACTGACACCAATGAGGCGCGCCCTGTGCTGCAACTCTGTGCACTGTCTT 380
Qy      301 CAGCAAAAGTAACATGCGCCACAGATCTGATGCTCTCAGTGGCCCTGGGGGAAACATAAT 360
Db      381 CAGCAAAAGTAACATGCGCCACAGATCTGATGCTCTCAGTGGCCCTGGGGGAAACATAAT 440
Qy      361 TACTGAGAGAACCCAGACAAACCGAGGCGGACCTGTGCTATGTGCAAGTGTGGCTTAAG 420
Db      441 TACTGAGAGAACCCAGACAAACCGAGGCGGACCTGTGCTATGTGCAAGTGTGGCTTAAG 500

```

```

Qy      421 CGGCTGTGCCAAGAGTGCATGTGTGATGACTGTGGGAGATGGAAGAAAGCCCTCTCTCT 480
Db      501 CGGCTGTGCCAAGAGTGCATGTGTGATGACTGTGGGAGATGGAAGAAAGCCCTCTCTCT 560
Qy      481 CCAGAAAGATTTAAATTTTCAGTGTGGCCAAAGAACTCTGAGGCCCGCTTTAAGATTAT 540
Db      561 CCAGAAAGATTTAAATTTTCAGTGTGGCCAAAGAACTCTGAGGCCCGCTTTAAGATTAT 620
Qy      541 GGGGGAGAAATTCACCAACCATCGAAGAACAGCCCTGTGTTGGGCCATTCAGAGGCGAC 600
Db      621 GGGGGAGAAATTCACCAACCATCGAAGAACAGCCCTGTGTTGGGCCATTCAGAGGCGAC 680
Qy      601 CGGGGGGGGCTGTGCTACCTGATGATGAGAGAGGCTGATCAAGCCCTGTGGGTGATC 660
Db      681 CGGGGGGGGCTGTGCTACCTGATGATGAGAGAGGCTGATCAAGCCCTGTGGGTGATC 740
Qy      661 AGGCCACACACTGCTTCAATTGATTAACCAAGAAAGAGAGACTACATCGTCTACCTGGGT 720
Db      741 AGGCCACACACTGCTTCAATTGATTAACCAAGAAAGAGAGACTACATCGTCTACCTGGGT 800
Qy      721 CGCTCAAGGCTTAATCTTCAACACGCAAGGGGAGATGAAGTTGAGGTGAAAACCTTAATC 780
Db      801 CGCTCAAGGCTTAATCTTCAACACGCAAGGGGAGATGAAGTTGAGGTGAAAACCTTAATC 860
Qy      781 CTACAGAAAGACTACAGCGCTGACAGCTTGTCTACACAAAGCAATTCGCTTGTGAGAG 840
Db      861 CTACAGAAAGACTACAGCGCTGACAGCTTGTCTACACAAAGCAATTCGCTTGTGAGAG 920
Qy      841 ATCCGTTTCAAGAGAGGAGGTGTGGCGAGCCATCCCGGACTTACAGACCATCTGCTGT 900
Db      921 ATCCGTTTCAAGAGAGGAGGTGTGGCGAGCCATCCCGGACTTACAGACCATCTGCTGT 980
Qy      901 CCTGTGATGTATACGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGCTTTGAAAA 960
Db      981 CCTGTGATGTATACGATCCCAAGTTTGGACAAAGCTGTGAGATCACTGCTTTGAAAA 1040
Qy      961 GAGAAATTCACGACATATCTTATCCGAGAGAGCTGAAAATGCTTTGAGAGCTGAT 1020
Db      1041 GAGAAATTCACGACATATCTTATCCGAGAGAGCTGAAAATGCTTTGAGAGCTGAT 1100
Qy      1021 TCCACCGGAGAGTGTGACGAGCCCACTACTACGCTCTGAAGTCAACCAAAATGCTG 1080
Db      1101 TCCACCGGAGAGTGTGACGAGCCCACTACTACGCTCTGAAGTCAACCAAAATGCTG 1160
Qy      1081 TGTGCTGTGACCCACAGTGTGAAAACAGATTTCTGTGCAAGGAGACTCAAGGGGAAACCCCTC 1140
Db      1161 TGTGCTGTGACCCACAGTGTGAAAACAGATTTCTGTGCAAGGAGACTCAAGGGGAAACCCCTC 1220
Qy      1141 GTCTGTTCCTCCAGGCGGCATGACTTTGACTGAAATTTGAGTGTGGGCGCGTGAATGT 1200
Db      1221 GTCTGTTCCTCCAGGCGGCATGACTTTGACTGAAATTTGAGTGTGGGCGCGTGAATGT 1280
Qy      1201 GCCCTGAAGAGACAGCCAGGCTGTACAGAGAGTCCACACTTTTACCTTGATCCGC 1260
Db      1281 GCCCTGAAGAGACAGCCAGGCTGTACAGAGAGTCCACACTTTTACCTTGATCCGC 1340
Qy      1261 AGTCACACCAAGAAAGAAATGGCTGTGGCCCTCTGA 1296
Db      1341 AGTCACACCAAGAAAGAAATGGCTGTGGCCCTCTGA 1376

```

Search completed: March 21, 2004, 21:00:58
Job time : 522 secs

THIS PAGE BLANK (USPTO)

FEATURES
source

Location/Qualifiers
1..870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6292200"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."|

ORIGIN

Query Match 63.5%; Score 822.6; DB 13; Length 870;
Best Local Similarity 98.7%; Pred. No. 1.9e-213;
Matches 850; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 75 TCAAGTTCATCGAAGCTGTGACTGTCTAAATGAGAGAACTGTCTCAACAAGTACTT 134
Db 1 TCAAGTTCATCGAAGCTGTGACTGTCTAAATGAGAGAACTGTCTCAACAAGTACTT 60

QY 135 CTCCAACATTCACTGGTGGCAATGCCCCCAAGAAATTTGGAGGGCAGCAGCTGTGAATAGA 194
Db 61 CTCCAACATTCACTGGTGGCAATGCCCCCAAGAAATTTGGAGGGCAGCAGCTGTGAATAGA 120

QY 195 TAACTCAAAAACTGTATGAGGGGAAATGCTCTTTACCGAGAAAGGCCAGCACTGA 254
Db 121 TAACTCAAAAACTGTATGAGGGGAAATGCTCTTTACCGAGAAAGGCCAGCACTGA 180

QY 255 CACATGGGCGGCGCTGCTGCTGCTGAACTGTGCACTGTCTTCAGCAAACTGACCA 314
Db 181 CACATGGGCGGCGCTGCTGCTGCTGAACTGTGCACTGTCTTCAGCAAACTGACCA 240

QY 315 TGCCACAGATCTGATCTCTTACAGTGGGCTGGGGAAACATTAATTCAGCAGAAACC 374
Db 241 TGCCACAGATCTGATCTCTTACAGTGGGCTGGGGAAACATTAATTCAGCAGAAACC 300

QY 375 AGACAAACCGAGGCGACCTGTGCTATGTGCAAGTGGGCTTAAAGCCGCTTGTCCAGA 434
Db 301 AGACAAACCGAGGCGACCTGTGCTATGTGCAAGTGGGCTTAAAGCCGCTTGTCCAGA 360

QY 435 GTGCATGTGTCAGACTGGCCGACATGGAATAAGCCCTCTCTCTCCAGAGATTAA 494
Db 361 GTGCATGTGTCAGACTGGCCGACATGGAATAAGCCCTCTCTCTCCAGAGATTAA 420

QY 495 ATTTCAGTGTGGCAAAAGACTCTGAGGCCCGCTTTAAGATTATTTGGGGAGAAATTC 554
Db 421 ATTTCAGTGTGGCAAAAGACTCTGAGGCCCGCTTTAAGATTATTTGGGGAGAAATTC 480

QY 555 CACCATTCGAGAACCAAGCCCTGTGTTGGGCTATCTACAGAGGCAACGGGGGGCTGTGT 614
Db 481 CACCATTCGAGAACCAAGCCCTGTGTTGGGCTATCTACAGAGGCAACGGGGGGCTGTGT 540

QY 615 CACCTAGTGTGTGAGGAGCCTCATACGCCCTTCTGGGTGATAGAGGCCACACACTG 674
Db 541 CACCTAGTGTGTGAGGAGCCTCATACGCCCTTCTGGGTGATAGAGGCCACACACTG 600

QY 675 CTTCATTGATTTCCCAAGAGAGAGACTACATCGTCTACCTGGTGGCTCAAGGCTTAA 734
Db 601 CTTCATTGATTTCCCAAGAGAGAGACTACATCGTCTACCTGGTGGCTCAAGGCTTAA 660

QY 735 CTTCACACGAGGAGGAGAGATGAAGTTTGAAGTGAAGAACTTAATCTTACACAGAGACT 794
Db 661 CTTCACACGAGGAGGAGAGATGAAGTTTGAAGTGAAGAACTTAATCTTACACAGAGACT 720

QY 795 CAGCGCTGACAGCGCTTGCTCAACAAGACATTTGCTGTGAAGATCCGTT-CCAAG 853
Db 720 CAGCGCTGACAGCGCTTGCTCAACAAGACATTTGCTGTGAAGATCCGTT-CCAAG 853

Db 721 CAGCGCTGACAGCGCTTGCTCAACAAGACATTTGCTGTGAAGATCCGTTCCAAG 780

QY 854 AGGGCAGAGTGTGGGAGCCATCCGGAGCTATACAGACATTCGCTG-CCCTCGATGAT 912
Db 781 AGGGCAGAGTGTGGGAGCCATCCGGAGCTATACAGACATTCGCTGCTCGATGAT 840

QY 913 AACGATCCCCAGTTTGGCACA 933
Db 841 AACGATCCCCAGTTTGGGCA 861

RESULT 2
CA489682 882 bp mRNA linear EST 14-NOV-2002
LOCUS AGENCOURT 10810373 MAFcl Homo sapiens cDNA clone IMAGE:6722162 5',
DEFINITION mRNA sequence.
ACCESSION CA489682
VERSION CA489682.1 GI:24952473
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC <http://mgs.mci.nih.gov/>.
1 (bases 1 to 882)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LMNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: LLM14284 row: h column: 02
High quality sequence start: 22
High quality sequence stop: 640.
Location/Qualifiers
1..882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722162"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HEM1, LNCaP"
/lab_host="EMD10B"
/clone_id="MAPCL"
/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dt. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan. Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

FEATURES
source

Query Match 62.5%; Score 809.6; DB 14; Length 882;
Best Local Similarity 96.6%; Pred. No. 7e-210;
Matches 847; Conservative 0; Mismatches 27; Indels 3; Gaps 2;

QY 127 AAGTACTTCCCAACATTCACTGTGCACTGCCCAAGAAATTTGGAGGGCAGCACTGT 186
Db 1 AAGGCTTGTAACNNCGGTCCGAATTCGCGGATTAAGAAATTCGAGGGCAGCACTGT 60

QY 187 GAATATAGTAAGTCAAAAACTGCTATGAGGGGAATGTCACTTTTACGAGGAAAGGCC 246
Db 61 GAATATAGTAAGTCAAAAACTGCTATGAGGGGAATGTCACTTTTACGAGGAAAGGCC 120

QY 247 AGCACTGACACATGGGCGGCGCTGCTGCTGGAATCTTGCACTGTCTTCAAGCAA 306
Db 247 AGCACTGACACATGGGCGGCGCTGCTGCTGGAATCTTGCACTGTCTTCAAGCAA 306

Db 121 AGCACTGACACCATGAGCCGCGCCCTGCTGCTGCACTGCTGCTTACGAA 180
 QY 307 AGTACCATGCCACGATCTGATGCTCTTCACTGGGCTGGGAAACATTAATTCTGC 366
 Db 181 AGGTACCAATGCCACGATCTGATGCTCTTCACTGGGCTGGGAAACATTAATTCTGC 240
 QY 367 AGGAACCCAGACACACGAGGAGGACCCCTGGTGTATGTGAGGTGGGCTTAAAGCGCTT 426
 Db 241 AGGAACCCAGACACACGAGGAGGACCCCTGGTGTATGTGAGGTGGGCTTAAAGCGCTT 300
 QY 427 GTTCAAGAGTGCATGTGTGATGATCTGCGAGATGAGAAAGAGCCCTCTCTCTCCAGAA 486
 Db 301 GTTCAAGAGTGCATGTGTGATGATCTGCGAGATGAGAAAGAGCCCTCTCTCTCCAGAA 360
 QY 487 GAATTAATTAATTTAGTGTGCGCAAAAGATCTGTAGAGCCCGCTTTAAGATTAATTTGGGGA 546
 Db 361 GAATTAATTAATTTAGTGTGCGCAAAAGATCTGTAGAGCCCGCTTTAAGATTAATTTGGGGA 420
 QY 547 GAATTCACCAACCATGAGAAACGAGCCCTGGTGTGCGGATCTTACAGAGAGGACCGGGGG 606
 Db 421 GAATTCACCAACCATGAGAAACGAGCCCTGGTGTGCGGATCTTACAGAGAGGACCGGGGG 480
 QY 607 GGCTCTGTCACTGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 666
 Db 481 GGCTCTGTCACTGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
 QY 667 ACACATCTGCTTCTTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 726
 Db 541 ACACATCTGCTTCTTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
 QY 727 AGGCTTAATCTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 786
 Db 601 AGGCTTAATCTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
 QY 787 AAGGACTACAGCGCTGACACGCTTGTCTACACAGACAGATTCCTTGTGAGATCCGT 846
 Db 661 AAGGACTACAGCGCTGACACGCTTGTCTACACAGACAGATTCCTTGTGAGATCCGT 720
 QY 847 TCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 906
 Db 721 TCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 QY 907 ATGTATTAAGAGTCCCAAGTTGGGACAGATCTGAGATCACTGG--CTTTGAGAAAGAG 963
 Db 781 ATGTATTAAGAGTCCCAAGTTGGGACAGATCTGAGATCACTGG--CTTTGAGAAAGAG 840
 QY 964 AATTCTACCGACTATCTATCTATCTGAGAGAGCTGAAAA 1000
 Db 841 AATTCTACCGACTATCTATCTATCTGAGAGAGCTGAAAA 877
 RESULT 3
 CFI32029 820 bp mRNA linear EST 05-AUG-2003
 LOCUS UI-HF-P00-aws-o-20-0-UI-r1 NIH_MGC_215 Homo sapiens cDNA clone
 DEFINITION IMAGE:30553987.5', mRNA sequence.
 ACCESSION CFI32029
 VERSION CFI32029.1 GI:33214878
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 820)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: Bento-Soares@uiowa.edu
 Tissue Procurement: Mary Hendrix
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/humanfl.html
 Seg primer: pYX-5.
 Location/Qualifiers
 1..820
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30553987"
 /issue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1ib="NIH_MGC_215"
 /notes="Organ: Lung; Vector: pYX-Asc; Site:1: EcoR I;
 Site:2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 61.9%; Score 802.2; DB 14; Length 820;
 Best Local Similarity 99.3%; Pred. No. 7e-208;
 Matches 815; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 218 GGAATGTCACCTTTACCGAGAAAGCCAGACATGACATGAGGCGGCGCTGCTGCTGC 277
 Db 1 GGAATGTCACCTTTACCGAGAAAGCCAGACATGACATGAGGCGGCGCTGCTGCTGC 60
 QY 278 CCTGGAATCTGCGCACTGTCTTACAGCAAGCATGATGCCACAGATCTGATGCTCTTC 337
 Db 61 CCTGGAATCTGCGCACTGTCTTACAGCAAGCATGATGCCACAGATCTGATGCTCTTC 120
 QY 338 AGCTGGGCTGGGAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 397
 Db 121 AGCTGGGCTGGGAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
 QY 398 GCTATGTGAGAGTGGGCTTAAAGCCGCTTGTCCAAAGTGCATGATGATGAGGAG 457
 Db 181 GCTATGTGAGAGTGGGCTTAAAGCCGCTTGTCCAAAGTGCATGATGATGAGGAG 240
 QY 458 ATGGAAGAAAGCCCTCTCTCTCCAGAGAAATTAATTAATTAATTAATTAATTAATTAATTA 517
 Db 241 ATGGAAGAAAGCCCTCTCTCTCCAGAGAAATTAATTAATTAATTAATTAATTAATTAATTA 300
 QY 518 TGAAGCCCGCTTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 577
 Db 301 TGAAGCCCGCTTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
 QY 578 TTGCGGCACTTACAGAGAGGACCCGGGGGGGCTCTGTCACTTACGCTGTGAGGAGGCC 637
 Db 361 TTGCGGCACTTACAGAGAGGACCCGGGGGGGCTCTGTCACTTACGCTGTGAGGAGGCC 420
 QY 638 TATACAGCCCTTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 697
 Db 421 TATACAGCCCTTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
 QY 698 AGGACTACATGCTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 757
 Db 481 AGGACTACATGCTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

QY 758 AGTTGAGTGGAAACCTAATCTACAGAGCACTACAGGCTGACAGCTTGTCTACC 817
 DB 541 AGTTGAGTGGAAACCTAATCTACAGAGCACTACAGGCTGACAGCTTGTCTACC 600
 QY 818 ACAACGACATTTGCTGCTGAGATCCGTTCCAGAGGCGAGGTGCGCAGCATCC 877
 DB 601 ACAACGACATTTGCTGCTGAGATCCGTTCCAGAGGCGAGGTGCGCAGCATCC 660
 QY 878 GGAATACACAGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 937
 DB 661 GGAATACACAGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 938 GTGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
 DB 721 GTGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
 QY 998 AATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
 DB 780 AATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820

RESULT 4
 BG741013 803 bp mRNA linear EST 15-MAY-2001
 LOCUS 602634740P1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779896 5',
 DEFINITION mRNA sequence.
 ACCESSION BG741013
 VERSION BG741013.1 GI:14051666
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 803)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph. D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: LIML0637 row: p column: 09
 High quality sequence stop: 784.
 Location/Qualifiers
 1. 803
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4779896"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Skn3"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NCI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
 Query Match 60.5%; Score 783.6; DB 12; Length 803;
 Best Local Similarity 99.4%; Pred. No. 8.5e-203;
 Matches 797; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 224 GTGACCTTTTACGAGAGAAAGCCAGACCTGACACATGCGCGCCCTGCTGCTGCTGCA 283
 DB 1 GTGACCTTTTACGAGAGAAAGCCAGACCTGACACATGCGCGCCCTGCTGCTGCTGCA 60

QY 284 ACTCTGCACTGCTGCTTTCAGCAAAAGTACATGCGCGAGAGATCTGATGCTTTCAGCTGG 343
 DB 61 ACTCTGCACTGCTGCTTTCAGCAAAAGTACATGCGCGAGAGATCTGATGCTTTCAGCTGG 120

QY 344 GCCTGGGGAACATATTAATCTGACAGAAACCCAGACCGAGGAGGACCTGTGTCTATG 403
 DB 121 GCCTGGGGAACATATTAATCTGACAGAAACCCAGACCGAGGAGGACCTGTGTCTATG 180
 QY 404 TGCAGGTGGGCTTAAAGCCGCTTGTTCAGAGATGATGATGATGATGATGATGATGATG 463
 DB 181 TGCAGGTGGGCTTAAAGCCGCTTGTTCAGAGATGATGATGATGATGATGATGATGATG 240
 QY 464 AAAAGCCCTCTCTCTCCAGAGAAATTAATTTCAAGTGGGCCAAAGACTGTGAGGC 523
 DB 241 AAAAGCCCTCTCTCTCCAGAGAAATTAATTTCAAGTGGGCCAAAGACTGTGAGGC 300
 QY 524 CCCGCTTTAAGATTAATGGGAGGAAATTCACCATCATGAGAACCGCCCTGTGTGGG 583
 DB 301 CCCGCTTTAAGATTAATGGGAGGAAATTCACCATCATGAGAACCGCCCTGTGTGGG 360
 QY 584 CCATCTACAGAGAGGACCGGGGGGCTGTGTACCTAGTGTGTGTGTGTGTGTGTGTGTGT 643
 DB 361 CCATCTACAGAGAGGACCGGGGGGCTGTGTACCTAGTGTGTGTGTGTGTGTGTGTGTGT 420
 QY 644 GCCCTTGTGAGTATGATGACGCGCACACATGCTTCAATGATTACCCAAAGAGAGACT 703
 DB 421 GCCCTTGTGAGTATGATGACGCGCACACATGCTTCAATGATTACCCAAAGAGAGACT 480
 QY 704 ACATGCTCTACTGCGGTGCTTCAAGGCTTAACTCCAAACCGAGGGAGATGAATTGG 763
 DB 481 ACATGCTCTACTGCGGTGCTTCAAGGCTTAACTCCAAACCGAGGGAGATGAATTGG 540
 QY 764 AGTGGAAAACTTAATCTTCAACAGAGACTACAGGCTGACAGCGCTTCTACCAACAAG 823
 DB 541 AGTGGAAAACTTATCTTCAACAGAGACTACAGGCTGACAGCGCTTCTACCAACAAG 600
 QY 824 ACATGCTCTTGTGAAATCCGTTCCAAAGAGGAGGTGTGCGCAGGCTCCGAGACTA 883
 DB 601 ACATGCTCTTGTGAAATCCGTTCCAAAGAGGAGGTGTGCGCAGGCTCCGAGACTA 660
 QY 884 TACACACCATCTGCTGCGCTCGATGTATTAACATATCCCAAG-TTTGGCAACAAGCTGTAG 942
 DB 661 TACACACCATCTGCTGCGCTCGATGTATTAACATATCCCAAG-TTTGGCAACAAGCTGTAG 720
 QY 943 ATCACTGCTTGTGAAAGAAATTTCCAGACTATCTATCCGAGCAGCTGAAATG 1002
 DB 721 ATCACTGCTTGTGAAAGAAATTTCCAGACTATCTATCCGAGCAGCTGAAATG 780

QY 1003 ACTGTTGTGAAGCTGATTTCC 1024
 DB 781 ACTGTTGTGAAGCTGATTTCC 802

RESULT 5
 BX360462 1200 bp mRNA linear EST 05-MAY-2003
 LOCUS BX360462 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS001070Y01 5-PRIME, mRNA sequence.
 ACCESSION BX360462
 VERSION BX360462.1 GI:30372436
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1200)
 Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5613.f For
 more information about this cluster, see

http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1070CB010P1&cluster=5613.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1070CB010P1.
 Location/Qualifiers

FEATURES

source

1..1200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1070YD01"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="Left strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 60.4%; Score 783.2; DB 13; Length 1200;
 Best Local Similarity 96.4%; Pred. No. 1.4e-202;
 Matches 812; Conservative 15; Mismatches 8; Indels 3; Gaps 3;

1 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGAGCGAATCCAAAGC 60
 169 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGAGCGAATCCAAAGC 228
 61 AGCAATGAATTCATCAATGTTCCATGCACTGCTGCTGCTTAATGAGAGAACTGTGT 120
 229 AGCAATGAATTCATCAATGTTCCATGCACTGCTGCTGCTTAATGAGAGAACTGTGT 288
 121 TCCAAAGAGTACTTCCCAATTCCTGCTGCACTGCCCAAGAAATTCGAGAGGCGAG 180
 289 TCCAAAGAGTACTTCCCAATTCCTGCTGCACTGCCCAAGAAATTCGAGAGGCGAG 348
 181 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGA 240
 349 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGA 408
 241 AAGGCGAGCACTGACACCATGAGGCGCGCCCTGCTGCTGCTGAACTGTGCTGCTCTT 300
 409 AAGGCGAGCACTGACACCATGAGGCGCGCCCTGCTGCTGCTGAACTGTGCTGCTCTT 468
 301 CAGCAACGTAACATCCCAACATCTGATGCTCTTCACTGAGGCGCTGGGAAACATTAAT 360
 469 CAGCAACGTAACATCCCAACATCTGATGCTCTTCACTGAGGCGCTGGGAAACATTAAT 528
 361 TACTGAGGAACCCAGAACCCGAGGCGAGCCCTGCTGCTATGAGGAGTGGGCTTAAG 420
 529 TACTGAGGAACCCAGAACCCGAGGCGAGCCCTGCTGCTATGAGGAGTGGGCTTAAG 587
 421 CCGCTTGTCCAGAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 588 CCGCTTGTCCAGAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 647
 481 CCGAAGAAATTAATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 648 CCGAAGAAATTAATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707
 541 GGGGGAGAAATTCACCAACATCGAGAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 708 GGGGGAGAAATTCACCAACATCGAGAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 767
 601 CCGGGGGGCTCTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 768 CCGGGGGGCTCTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827
 661 AGGCGCAACACGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 828 AGGCGCAACACGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
 721 CCGTCAAGGCTTAATCTCAACAGCAAGGGAGATGATGATGATGATGATGATGATGATGAT 780

Db 888 CGCTCAAGGCTTAATCTCAACAGCAAGGGAGATGATGATGATGATGATGATGATGATGAT 946
 781 CTACACAGAGACTACAGAGGCTGACAGGCTTGTCTCAACACAGCAATTCGCTTCTGA 838
 947 MTACAAAGAGACTACAGAGGCTGACAGGCTTGTCTCAACACAGCAATTCGCTTCTGA 1003

RESULT 6
 LOCUS BUI90752 1003 bp mRNA linear EST 04-SEP-2002
 DEFINITION AGENCOURT_6925417 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5952454
 5', mRNA sequence.
 ACCESSION BUI90752
 VERSION BUI90752.1 GI:22704736
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1003)
 NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHC2140 row: h column: 23
 High quality sequence stop: 608.

FEATURES

source

1..1003
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5952454"
 /issue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 110"
 /note="Organ: pancreas; Vector: pOT87; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.3%; Score 781; DB 13; Length 1003;
 Best Local Similarity 98.4%; Pred. No. 5e-202;
 Matches 809; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

1 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGAGCGAATCCAAAGC 60
 73 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGAGCGAATCCAAAGC 132
 61 AGCAATGAATTCATCAATGTTCCATGCACTGCTGCTGCTTAATGAGAGAACTGTGT 120
 133 AGCAATGAATTCATCAATGTTCCATGCACTGCTGCTGCTTAATGAGAGAACTGTGT 192
 121 TCCAAAGAGTACTTCCCAATTCCTGCTGCACTGCCCAAGAAATTCGAGAGGCGAG 180
 193 TCCAAAGAGTACTTCCCAATTCCTGCTGCACTGCCCAAGAAATTCGAGAGGCGAG 252
 181 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTTACGAGGA 240
 253 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTTACGAGGA 312

OY		241	AAGGCGACGACTGCACCAATGGGCGGACCCTGGCTBGCCTGGAACCTTCGCACTGTCCCTT	300
Db		313	AAGGCCAGACTGACACCATGGCCGGGCCCCCTGCTCCTCCCTGGAACTCTGCATGTGCTTT	372
OY		301	CAGCAAACGTACCATGCCACAAGATCTGA TGCTCTTCACTGGGCTTGGGAAAACATAAT	360
Db		373	CAGCAAACTTACCATCCCAAGATCTGA TGTGCTCTTCACTGGGCTTGGGAAAACATAAT	432
OY		361	TACTGAGAACCACCAACAACCGAGGCGAACCTTGTTGTATGTGAGGTGGGCTTAAG	420
Db		433	TACTGAGAACCCCAACAACCGAGGCGAACCTTGTTGTATGTGAGGTGGGCTTAAG	492
OY		421	CCGCTTTGTCAAAGATGATGTCATGTCATGTCGGCAGATGAAAAAAGCCCTCTCTCT	480
Db		493	CCGCTTTGTCAAAGATGATGTCATGTCATGTCGGCAGATGAAAAAAGCCCTCTCTCT	552
OY		481	CCAGAAAGAAATTAAAAATTTCAGTGTGGCCAAAAGACTCTGAGGCCCGCTTTAAGATTAT	540
Db		553	CCAGAAAGAAATTAAAAATTTCAGTGTGGCCAAAAGACTCTGAGGCCCGCTTTAAGATTAT	612
OY		541	GGGGGAGAAATTGACCAACCATCGAGAA CAGCCCTGGTTTGGGGCATCTACAGAGGACAC	600
Db		613	GGGGGAGAAATTGACCAACCATCGAGAA CAGCCCTGGTTTGGGGCATCTACAGAGGACAC	672
OY		601	CGGGGGGGCTCTGTGCACTTACGTGTGTGAGGACAGCTTCATCAGCCCTTGTGAGTGCATC	660
Db		673	CGGGGGGGCTCTGTGCACTTACGTGTGTGAGGACAGCTTCATCAGCCCTTGTGAGTGCATC	732
OY		661	AGCGGCACACACTGTGCTTCAATTGATTA CCCAAAGAAAGAGAGACTACATGCTCTACCTGGCT	720
Db		733	AGCGGCACACACTGTGCTTCAATTGATTA CCC-AAGAAAGAGAGACTACATGCTCTACCTGGCT	791
OY		721	CGCTCAAGGCTTAACCTCCAACA CGCAAGGGAGATGAATTGAGTGGAAAACTTAATC	780
Db		792	CGCTCAAGGCTTAACCTCCAACA CGCAAGGGAGATGAATTGAGTGGAAAACTTAATC	851
OY		781	CTACACAAAGAGACTACAGC-GCTGACAC GCTTGCTCTACCAACA 821	
Db		852	CTACACAAAGAGATCACCGGCGCTGA AAAACGTTGCTTCCACCA 893	
RESULT 7				
CFI32052				
LOCUS		826 bp	mRNA	linear EST 05-AUG-2003
DEFINITION		UT-HF-PQ0-aws-d-23-0-UT.r1 NHE_MGC_215 Homo sapiens cDNA clone IMAGE:30553726 5', mRNA sequence.		
ACCESSION		CFI32052		
VERSION		CFI32052.1 GI:3214924		
KEYWORDS		EST,		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 826) Bonaldo,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)		
AUTHORS		JOURNAL		
TITLE		MEDLINE		
PUBMED		COMMENT		
		8889548		
		Contact: Soares, MB		
		Coordinated Laboratory for Computational Genomics		
		University of Iowa		
		375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA		
		Tel: 319 335 8250		
		Fax: 319 335 9565		
		Email: bento-soares@uiowa.edu		
		Tissue Procurement: Mary Hendrix		
		cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
		cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
		Clone Sequencing by: Dr. M. Bento Soares, University of Iowa		
		CDNA Distribution: Distribution information can be found at		
		http://genome.uiowa.edu/distribution/humanmtl.html		

FEATURES		Seq primer: pYX-5.
SOURCE		location/Qualifiers
1. 826		
/organism="Homo sapiens"		
/mol_type="mRNA"		
/db_xref="taxon:9606"		
/clone="IMAGE:30553726"		
/lab_host="Chondrosarcoma Lung Metastasis cell lines"		
/feature_type="PHI08 (T1 phage resistant)"		
/olone_11b="NH1_MGC_215"		
/note="Organ: Lung; Vector: pYX-Asc; Site: 1: EcoR I; Site: 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATTAAGGCCA. Tissue was provided by Mary Hendrix."		
ORIGIN		
Query Match	60.2%; Score 779.8; DB 14; Length 826;	
Best Local Similarity	99.3%; Pred. No. 9.4e-202;	
Matches 814; Conservative	0; Mismatches 3; Indels 3; Gaps 3	
OY	225 TCACCTTTTACGAGGAAAGGCGACACATGCGCGGCGCTGCGCTCGCTGAA	284
Db	1 TCACCTTTTACGAGGAAAGGCGACACATGCGCGGCGCTGCGCTCGCTGAA	59
OY	285 CTCGCGACTGCTCTTCAGCAACGTCACATGCCACAGATCTGATCTTTCAGTGG	344
Db	60 CTCGCGACTGCTCTTCAGCAACGTCACATGCCACAGATCTGATCTTTCAGTGG	119
OY	345 CCTGGGGGAAACATAATTCTGCAGAAACCCAGAACACCGAGGCGACCTTGCTGATGT	404
Db	120 CCTGGGGGAAACATAATTCTGCAGAAACCCAGAACACCGAGGCGACCTTGCTGATGT	179
OY	405 GGAGGTGGGCGCTTAAAGCGCGCTTGTCAGAGTGCATGTGTGATGATCTGGCAGATGAA	464
Db	180 GGAGGTGGGCGCTTAAAGCGCGCTTGTCAGAGTGCATGTGTGATGATCTGGCAGATGAA	239
OY	465 AAAGCCCTCTCTCTCTCCAGAAATTTAAATTTCACTGTGGCGCAAAAGACTCTGAGGCC	524
Db	240 AAAGCCCTCTCTCTCTCCAGAAATTTAAATTTCACTGTGGCGCAAAAGACTCTGAGGCC	299
OY	525 CGGCTTTAAGATTATTTGGGGGAGAAATTCACACCAATTCGAGAACCGAGCCCTGGTTGGCGG	584
Db	300 CGGCTTTAAGATTATTTGGGGGAGAAATTCACACCAATTCGAGAACCGAGCCCTGGTTGGCGG	359
OY	585 CATCTACAGAGGACCGGGGGGCGCTCTGCATCTACGTGTGTGGAGGACGCTCATGAG	644
Db	360 CATCTACAGAGGACCGGGGGGCGCTCTGCATCTACGTGTGTGGAGGAGGCTCATGAG	419
OY	645 CCCCTTGCTGGGTGATCAGCGGCACACATGCTTCATTGATTAACCCAAAGAGAGAGACTA	704
Db	420 CCCCTTGCTGGGTGATCAGCGGCACACATGCTTCATTGATTAACCCAAAGAGAGAGACTA	479
OY	705 CATGCTTAACCTGGGTGGCTCAAGGCTTAACTCCACACAGCAAGGGGAGATGAAGTTGA	764
Db	480 CATGCTTAACCTGGGTGGCTCAAGGCTTAACTCCACACAGCAAGGGGAGATGAAGTTGA	539
OY	765 GGTGAAAACTTAATCTTACACAGAGCTACAGCGCTGACAGCTTGTCTCCACACAGA	824
Db	540 GGTGAAAACTTAATCTTACACAGAGCTACAGCGCTGACAGCTTGTCTCCACACAGA	599
OY	825 CATGGCTTGTGTAAGATCCGTTCCAAAGAGGGCAGGTGTGGGCGCATCCCGGACTAT	884
Db	600 CATGGCTTGTGTAAGATCCGTTCCAAAGAGGGCAGGTGTGGGCGCATCCCGGACTAT	659
OY	885 ACAGACATCTGCGCTGCGCTCGATGATTAACGATCCCGAGTTGGCAACAAGTGAAGAT	944

Db 660 ACAGACCATCTGCTCCCTCGATGTATAAGATCCCAAGTTGGACAGCTGTGAGAT 719

QY 945 CACTGCTTTGGAAGAAGATTCTACCGACTATCTTATCCGAGAGAGTGAATAATGAC 1004

Db 720 CACTGCTTTGGAAGAAG-ATTCTACCGACTATCTTATCCGAGAGAGTGAATAATGAC 778

QY 1005 TGTGTGAAGCTGATTTCCACCGGAGAGTGTGACAGAGCC 1044

Db 779 TGTGTGAAGCTGATTTCCCA-CGGAGTGTGACAGAGCC 817

RESULT 8
CF132539 827 bp mRNA linear EST 05-AUG-2003
LOCUS UI-HF-F00-awg-p-06-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
DEFINITION IMAGE:30559899 5', mRNA sequence.

ACCESSION CF132539 GI:33215909
VERSION CF132539
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 827)
AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Mary Hendrix
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pyx-5

FEATURES
source
Location/Qualifiers
1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30559899"
/issue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="PH10B (T1 phage resistant)"
/clone_1id="NIH_MGC_215"
/note="Organ: Lung; Vector: pYX-Absc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. ligated with Ecor I adaptor, digested with Not I and then cloned directionally into pYX-Absc vector. The library tag sequence located between the Not I site and the polyA tail is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN
Query Match 60.0%; Score 777.6; DB 14; Length 827;
Best Local Similarity 98.3%; Pred. No. 3.8e-201;
Matches 816; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 225 TCACCTTTTACGAGAAAGCCAGCACTGACACCATGCGCGCCCTGCTGCTGCGAA 284

Db 1 TCACCTTTTACGAGAAAGG-CAGCACTGACACCATGCGCGCCCTGCTGCTGCGAA 59

QY 285 CTCGCACTGCTCTCAGCAAGCTACATGCCACAGATCTGATCTCTTACGTGGG 344

Db 60 CTCGCACTGCTCTCAGCAAGCTACATGCCACAGATCTGATCTCTTACGTGGG 119

QY 345 CCTGGGAAAACATATTAATGACAGAACCAACCGAGGCGACCTGTGTATGT 404

Db 120 CCTGGGAAAACATATTAATGACAGAACCAACCGAGGCGACCTGTGTATGT 179

QY 405 GCAAGTGGGCTTAAGCCGCTTGTCCAGAGTCAATGTGATGATCTGCCAATGAAA 464

Db 180 GCAAGTGGGCTTAAGCCGCTTGTCCAGAGTCAATGTGATGATCTGCCAATGAAA 239

QY 465 AAGAGCCCTCTCTCCCAAGAAATTAATTCAGTGGGCAAAAGACTGAGGCC 524

Db 240 AAGAGCCCTCTCTCTCCCAAGAAATTAATTCAGTGGGCAAAAGACTGAGGCC 299

QY 525 CCGCTTTAAGATTATGGGGGAGAAATTCACCAACATCGAAGAACCGCTGTTGCGG 584

Db 300 CCGCTTTAAGATTATGGGGGAGAAATTCACCAACATCGAAGAACCGCTGTTGCGG 359

QY 585 CATCTACAGAGGACACCGGGGGGCTGTGTCACTACGTGTGTGAGAGGCTCATG 644

Db 360 CATCTACAGAGGACACCGGGGGGCTGTGTCACTACGTGTGTGAGAGGCTCATG 419

QY 645 CCGTGTGGGTGATGAGGCGCACACATGCTTCAATTAACCAAGAGAGGAGCTA 704

Db 420 CCGTGTGGGTGATGAGGCGCACACATGCTTCAATTAACCAAGAGAGGAGCTA 479

QY 705 CATGCTACCTGGGTGCTCAAGGCTTAATCTCAACAGCAAGGAGATGAAGTTGA 764

Db 480 CATGCTACCTGGGTGCTCAAGGCTTAATCTCAACAGCAAGGAGATGAAGTTGA 539

QY 765 GGTGAAAACCTATCTTACACAGGACTACAGGCTGACAGGCTTCTTACCAACGA 824

Db 540 GGTGAAAACCTATCTTACACAGGACTACAGGCTGACAGGCTTCTTACCAACGA 599

QY 825 CATTGCTGTGGAATATCGTTCCAGAGAGGAGAGTGTGCGCAACATCCGGACTAT 884

Db 600 CATTGCTGTGGAATATCGTTCCAGAGAGGAGAGTGTGCGCAACATCCGGACTAT 659

QY 885 ACAGACCATCTGCTCCCTCGATGTATTAACGATCCCAAGTTGGACAGCTGTGAGAT 944

Db 660 ACAGACCATCTGCTCCCTCGAT-TAATACATCCCAAGTTGGACAGCTGTGAGAT 718

QY 945 CACTGCTTTGGAAGAAGATTCTACCGACTATCTTATCCGAGAGCTGAATAATGAC 1004

Db 719 CACTGCTTTGGAAGAAG-ATTCTACCGACTATCTTATCCGAGAGCTGAATAATGAC 777

QY 1005 TGTGTGAAGCTGATTTCCACCGGAGAGTGTGACAGAGCCCACTACTAGG 1054

Db 778 TGTGTGAAGCTGATTTCCACCGGAGAGTGTGACAGAGCCCACTACTAGG 827

RESULT 9
CF131558 814 bp mRNA linear EST 05-AUG-2003
LOCUS UI-HF-F00-awg-p-06-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
DEFINITION IMAGE:30559899 5', mRNA sequence.

ACCESSION CF131558
VERSION CF131558.1 GI:33213934
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 814)
AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu

Tissue Procurement: Mary Hendrix
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/humanf1.html>
Seq primer: PYX-5.

FEATURES

Source

Location/Qualifiers

1..814
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30558989"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH_MGC_215"
/note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 59.9%; Score 776.8; DB 14; Length 814;
Best Local Similarity 99.3%; Pred. No. 6.2e-201;
Matches 811; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

224 GTCACTTTTACCGGAAAGCCGACATGACATGAGCCGCTGCTGCTGCTGCA 283
1 GTCACTTTTACCGGAAAGG-CAGCACTACACATGAGCCGCTGCTGCTGCA 59

284 ACTGTGCACATGCTCTTCAAGCAACGATGATGATGATGATGATGATGATG 343
60 ACTGTGCACATGCTCTTCAAGCAACGATGATGATGATGATGATGATGATG 119

344 GCGTGGGAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403
120 GCGTGGGAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 179

404 TGCAGTGGGCTTAAAGCCGCTTGTCCAGATGATGATGATGATGATGATGAT 463
180 TGCAGTGGGCTTAAAGCCGCTTGTCCAGATGATGATGATGATGATGATGAT 239

464 AAAAGCCCTCTCTCTCTCAAGAAATTAATTAATTAATTAATTAATTAATTA 523
240 AAAAGCCCTCTCTCTCTCAAGAAATTAATTAATTAATTAATTAATTAATTA 299

524 CCGGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 583
300 CCGGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 359

584 CCACTTACAGAGGACACCGGGGGGCTGTGTACACTACGATGTGTGAGGACGCTCAT 643
360 CCACTTACAGAGGACACCGGGGGGCTGTGTACACTACGATGTGTGAGGACGCTCAT 419

644 GCGCTTGT 703
420 GCGCTTGT 479

704 ACATCGTCTACCTGGGTGCTCAAGCTTAACCTCAACACGCAAGGGGAGATGAAGTTTG 763

Db

480 ACATCGTCTACCTGGGTGCTCAAGCTTAACCTCAACACGCAAGGGGAGATGAAGTTTG 539

Qy

764 AGGTGAAAACCTTAATCTTACACAGAGATTAACAGGCTGACACGCTTGTCAACCAAG 823

Db

540 AGGTGAAAACCTTAATCTTACACAGAGATTAACAGGCTGACACGCTTGTCAACCAAG 599

Qy

824 ACATGCTTGTGTAAGATCCGCTTCAAGAGAGGAGGTGTGCGACATCCGAGACTA 883

Db

600 ACATGCTTGTGTAAGATCCGCTTCAAGAGAGGAGGTGTGCGACATCCGAGACTA 659

Qy

884 TACAGACCATCTGCTGCTGCTTGTATATACGATCCCAAGTTTGGCAACAGTTGTAGA 943

Db

660 TACAGACCATCTGCTGCTGCTTGTATATACGATCCCAAGTTTGGCAACAGTTGTAGA 719

Qy

944 TCACCTGCTTTGAAAAGAAATTTACCGACTATCTATCCGAGACGCTGAAAATGA 1003

Db

720 TCACCTGCTTTGAAAAGAAATTTACCGACTATCTATCCGAGACGCTGAAAATGA 778

Qy

1004 CTGTTGTGAAGCTGATTTCCACCGGAGGTGCACA 1040

Db

779 CTGTTGTGAAGCTGATTT-CCACCGGAGGTGCACA 814

RESULT 10
LOCUS
BM920114
DEFINITION
AGENCOURT 6706522 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5749987
5', mRNA sequence.
ACCESSION
BM920114
VERSION
BM920114.1 GI:19370493
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1067)
NIH-MGC <http://mgi.nci.nih.gov/>.
JOURNAL
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: L1AM12780 row: d column: 20
High quality sequence stop: 754.
Location/Qualifiers
1..1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5749987"
/lab_host="DH10B"
/clone_id="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector: pcwv-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 59.6%; Score 772.4; DB 12; Length 1067;
Best Local Similarity 93.5%; Pred. No. 1.2e-199;
Matches 880; Conservative 0; Mismatches 48; Indels 13; Gaps 7;

QY 1 ATGAGAGCCCTGTGGGGGCTTCTCTGTGCTGCTGCTGAGGCACTCCAAAGC 60
 DB 127 ATGAGAGCCCTGTGGGGGCTTCTCTGTGCTGCTGAGGCACTCCAAAGC 186
 QY 61 AGCAATGAATTCATCATCAAGTTCATGCACTGCTGTAATGAGAGACATGTGTG 120
 DB 187 AGCAATGAATTCATCATCAAGTTCATGCACTGCTGTAATGAGAGACATGTGTG 246
 QY 121 TCCAAACAAGTACTTCTCCAACTTCACTGTGCACTGCCCAAAGAAATTCGAGAGG 180
 DB 247 TCCAAACAAGTACTTCTCCAACTTCACTGTGCACTGCCCAAAGAAATTCGAGAGG 306
 QY 181 CACTGTGAATATAGTAAAGTAAACCTGTATAGAGGGAAATGCTACTTTTACGAGAG 240
 DB 307 CACTGTGAATATAGTAAAGTAAACCTGTATAGAGGGAAATGCTACTTTTACGAGAG 366
 QY 241 AAGGCCAGACTGACACATGAGGCGGAGCCCTGCTGCAACTCTGCACTGTCTT 300
 DB 367 AAGGCCAGACTGACACATGAGGCGGAGCCCTGCTGCAACTCTGCACTGTCTT 426
 QY 301 CAGCAAAAGTACATCCATCCCAAGATCTGATCTTCACTGAGGCTGAGGAAACATTA 360
 DB 427 CAGCAAAAGTACATCCATCCCAAGATCTGATCTTCACTGAGGCTGAGGAAACATTA 486
 QY 361 TACTGAGAGAACCCAGACCAACCGAGAGGCAACCTGTGCTATGTGAGAGTGGCCCTTAAG 420
 DB 487 TACTGAGAGAACCCAGACCAACCGAGAGGCAACCTGTGCTATGTGAGAGTGGCCCTTAAG 546
 QY 421 CGGCTGTCCAAAGAGTGCATGTGTCATGACTGCGCAGATGGAATAAAGCCCTCTCTCT 480
 DB 547 CGGCTGTCCAAAGAGTGCATGTGTCATGACTGCGCAGATGGAATAAAGCCCTCTCTCT 606
 QY 481 CCAGAAAGATTTAAATTTCAAGTGTGSCCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
 DB 607 CCAGAAAGATTTAAATTTCAAGTGTGSCCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 666
 QY 541 GGGGGAGATTTACCAACCAATCCAGAACCAAGCCCTGTGTTGGGCCATCTACAGAGGCA 600
 DB 667 GGGGGAGATTTACCAACCAATCCAGAACCAAGCCCTGTGTTGGGCCATCTACAGAGGCA 726
 QY 601 CGGGGGGGGCTGTGTACCTACGTGTGAGAGGAGGCTCTACAGCCCTTGTGGTGTATC 660
 DB 727 CGGGGGGGGCTGTGTACCTACGTGTGAGAGGAGGCTCTACAGCCCTTGTGGTGTATC 786
 QY 661 AGCGGCACACACTGCTTCAATTGATTAACCAAGAGAGAGACTACATCGTCTA-CCTGGG 719
 DB 787 AGCGGCACACACTGCTTCAATTGATTAACCAAGAGAGAGACTACATCGTCTA-CCTGGG 846
 QY 720 TGGCTCAAGGCTTAATCTCCAAACCGCAAGG--GAGATGAAGTTGAGGTGAAAACT 776
 DB 847 TGGCTCAAGGCTTAATCTCCAAACCGCAAGG--GAGATGAAGTTGAGGTGAAAACT 906
 QY 777 AATCTCAACAAGAGTACAGGCG-TCACACGCTGTGTCAACCAAGACATTTGCTT-- 833
 DB 907 CATCTCAACAAGAGTACAGGCGCTTGAACAGCTTGTCTCAACCAAGACATTTGCTT 966
 QY 834 -GCTGAAGATCCGTTTCAAGAGAGGAGGTGT--GCGAGCCATCCCGACTATAC-A 888
 DB 967 CTGGAAGATCCGTTTCAAGAGAGGAGGTGTGTCANCAATCCCGAGATATACAA 1026
 QY 889 ACCATCTGCTGCTTCCCT-CCATGATATACGATCCCGAGTTT 928
 DB 1027 ANCATCTGCTGCTTCCCGAGAGATTAACGATCCCGCATTTG 1067

RESULT 11

LOCUS BU501434 925 bp mRNA linear EST 12-SEP-2002
 DEFINITION AGENCOURT 8877614 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6318152
 5' mRNA sequence.
 ACCESSION BU501434
 VERSION BU501434.1 GI:22803985

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 925)
 NIH-MGC <http://imgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LCM2530 row: 0 column: 17
 High quality sequence stop: 663.
 Location/Qualifiers
 1..925
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6318152"
 /feature_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOT7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. |"

ORIGIN
 Query Match 59.5%; Score 770.6; DB 13; Length 925;
 Best Local Similarity 98.3%; Pred. No. 3.3e-199;
 Matches 800; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

QY 1 ATGAGAGCCCTGTGGGGGCTTCTCTGTGCTGCTGAGGCACTCCAAAGC 60
 DB 104 ATGAGAGCCCTGTGGGGGCTTCTCTGTGCTGCTGAGGCACTCCAAAGC 163
 QY 61 AGCAATGAATTCATCATCAAGTTCATGCACTGCTGTAATGAGAGACATGTGTG 120
 DB 164 AGCAATGAATTCATCATCAAGTTCATGCACTGCTGTAATGAGAGACATGTGTG 223
 QY 121 TCCAAACAAGTACTTCTCCAACTTCACTGTGCACTGCCCAAAGAAATTCGAGAGG 180
 DB 224 TCCAAACAAGTACTTCTCCAACTTCACTGTGCACTGCCCAAAGAAATTCGAGAGG 283
 QY 181 CACTGTGAATATAGTAAAGTAAACCTGTATAGAGGGAAATGCTACTTTTACGAGAG 240
 DB 284 CACTGTGAATATAGTAAAGTAAACCTGTATAGAGGGAAATGCTACTTTTACGAGAG 343
 QY 241 AAGGCCAGACTGACACATGAGGCGGAGCCCTGCTGCAACTCTGCACTGTCTT 300
 DB 344 AAGGCCAGACTGACACATGAGGCGGAGCCCTGCTGCAACTCTGCACTGTCTT 403
 QY 301 CAGCAAAAGTACATCCATCCCAAGATCTGATCTTCACTGAGGCTGAGGAAACATTA 360
 DB 404 CAGCAAAAGTACATCCATCCCAAGATCTGATCTTCACTGAGGCTGAGGAAACATTA 463
 QY 361 TACTGAGAGAACCCAGACCAACCGAGAGGCAACCTGTGCTATGTGAGAGTGGCCCTTAAG 420
 DB 464 TACTGAGAGAACCCAGACCAACCGAGAGGCAACCTGTGCTATGTGAGAGTGGCCCTTAAG 523
 QY 421 CGGCTGTCCAAAGAGTGCATGTGTCATGACTGCGCAGATGGAATAAAGCCCTCTCTCT 480

Db 524 CCGCTTGTCCAGAGTGCATGTCATGCTGCGCAGATGCGAAAAAGCCCTCTCTCT 583
 QY 441 CCAGAGAAATTAATTAATTTAGTGTGCGCCAAAAGACTGTAGAGCCCGCTTTAGATTAT 540
 Db 584 CCAGAGAAATTAATTAATTTAGTGTGCGCCAAAAGACTGTAGAGCCCGCTTTAGATTAT 643
 QY 541 GGGGGAGAAATTCACCATCGAGAACGAGCCCTGTGCGGCATCTTACAGAGGAC 600
 Db 644 GGGGGAGAAATTCACCATCGAGAACGAGCCCTGTGCGGCATCTTACAGAGGAC 703
 QY 601 GGGGGAGGCTCTGTTCACCTAGCTGTGTGTGAGAGCAGCCCTCATCAGCCCTGTGAGTATC 660
 Db 704 GGGGGAGGCTCTGTTCACCTAGCTGTGTGTGAGAGCAGCCCTCATCAGCCCTGTGAGTATC 763
 QY 661 AGCGGCACACACTGTTTATTGATTATCCCAAGAGAGAGACTACATGCTTACCTGGT 720
 Db 764 AGCGGCACACACTGTTTATTGATTATCCCAAGAGAGAGACTACATGCTTACCTGGT 823
 QY 721 CGCTCAAGGCTTAACTCCACAGCAGGAGAGATGAGTTG--AGGTGAGAAACCTTA 778
 Db 824 CGCTCAAGGCTTAACTCCACAGCAGGAGAGATGAGTTGAGAGTGGAGAAACCTTA 883
 QY 779 TCCTACACAGG--ACTACAGCGCTGACACGCTT 810
 Db 884 TCCTACACAGGAGACTACAGCGCTGACACGCTT 917

RESULT 12

CF131569

LOCUS 778 bp mRNA linear EST 05-AUG-2003
 DEFINITION UI-HF-F00-awr-a-09-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
 IMAGE:30555176 5', mRNA sequence.

ACCESSION

CF131569

VERSION 1
 KEYWORDS GI:33213957
 EST.

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: Bento-soares@uiowa.edu
 Tissue Procurement: Mary Hendrix
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/humanfl.html>
 Seg primer: pYX-5.

FEATURES
 source Location/Qualifiers

1..778
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30555176"
 /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_MGC_215"
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose

ORIGIN

Query Match 59.4%; Score 769.4; DB 14; Length 778;
 Best Local Similarity 99.6%; Pred. No. 6,4e-195;
 Matches 770; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with Not I and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GATTAAGGCCA. Tissue was provided by Mary Hendrix.

QY 48 CGACTCCAAAGGACAGATGAATTCATCAAGTTCATGAACTGTGACTGTCTAAATGG 107
 Db 1 CGACTCCAAAGGACAGATGAATTCATCAAGTTCATGAACTGTGACTGTCTAAATGG 60
 QY 108 AGGAACATGTGTGTCCAAAGTACTTCTCCAACTTCACTGTGTGAATCTGCCAAGAA 167
 Db 61 AGGAACATGTGTGTCCAAAGTACTTCTCCAACTTCACTGTGTGAATCTGCCAAGAA 120
 QY 168 ATTCGAGAGGACAGACTGTGAAATAGATAGCAAAAACTGTATGAGGGAAATGTCA 227
 Db 121 ATTCGAGAGGACAGACTGTGAAATAGATAGCAAAAACTGTATGAGGGAAATGTCA 180
 QY 228 CTTTACCGAGAAAGGCCAGACTGACACATGAGGCGGCTTGCCTGCTGAACTC 287
 Db 181 CTTTACCGAGAAAGGCCAGACTGACACATGAGGCGGCTTGCCTGCTGAACTC 240
 QY 288 TGCACATGCTCTTCCAGAAAGTACATGAGTCCAGATCTGATCTGTGAGTGGGCT 347
 Db 241 TGCACATGCTCTTCCAGAAAGTACATGAGTCCAGATCTGATCTGTGAGTGGGCT 300
 QY 348 GGGGAAACATATTAATCTCAGAAACCCAGAACCCGAGAGGAGGAGCTGTGCTATGTGA 407
 Db 301 GGGGAAACATATTAATCTCAGAAACCCAGAACCCGAGAGGAGGAGCTGTGCTATGTGA 360
 QY 408 GGTGGGCTTAAAGCCGCTTGTCCAAAGTGTGATGATGCTGCGGAGTGGAAAAA 467
 Db 361 GGTGGGCTTAAAGCCGCTTGTCCAAAGTGTGATGATGCTGCGGAGTGGAAAAA 420
 QY 468 GCCCTCCTCTCCTCAGAAATTAATTTCAAGTGTGAGGAGGAGGAGGAGGAGGAGG 527
 Db 421 GCCCTCCTCTCCTCAGAAATTAATTTCAAGTGTGAGGAGGAGGAGGAGGAGGAGG 480
 QY 528 CTTTAAATTAATTTGAGGAGAAATTCACCAATGAGAACCAAGGCTGTGTTGGGAGCAT 587
 Db 481 CTTTAAATTAATTTGAGGAGAAATTCACCAATGAGAACCAAGGCTGTGTTGGGAGCAT 540
 QY 588 CTACAGAGAGGACCGGGGGGCTGTGTCACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 647
 Db 541 CTACAGAGAGGACCGGGGGGCTGTGTCACTAGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 QY 648 TTGCTGAGTGTATGAGCGGCACACACTGCTTATTGATTACCAAGAGAGGAGACTACAT 707
 Db 601 TTGCTGAGTGTATGAGCGGCACACACTGCTTATTGATTACCAAGAGAGGAGACTACAT 660
 QY 708 CGTCTACTGT 767
 Db 661 CGTCTACTGT 720
 QY 768 GGAAGAACTAATCTTACACAGAGACTACAGGCTGACAGCTTGTGTGTGTGTGTGTGTGT 820
 Db 721 GGAAGAACTAATCTTACACAGAGACTACAGGCTGACAGCTTGTGTGTGTGTGTGTGTGT 773

RESULT 13

CF132245

LOCUS 773 bp mRNA linear EST 05-AUG-2003
 DEFINITION UI-HF-F00-awr-a-14-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
 IMAGE:30555197 5', mRNA sequence.

ACCESSION

CF132245

VERSION

CF132245.1 GI:33215316

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Euxariyote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
AUTHORS	1 (bases 1 to 773) Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
PUBMED	8889548
COMMENT	Contact: Soares, MB


```
Db      541 CACCACCAAATGCTGTGTGCTGTGACCCACAGTGAATAACAGATTCTGCCAGGAGA 600
QY      1125 CTGAGGGGGACCCCTGCTGTGTTCCCTCCAAAGCCCGATGACTTTGACTGGAATTGTGAG 1184
Db      601 CTCAGGGGGACCCCTGCTGTGTTCCCTCCAAAGCCCGATGACTTTGACTGGAATTGTGAG 660
QY      1185 CTGGGGCCGTGATGTGCTGTGAAGGACAAGCCAGGCGCTTACACGAGAGTTCACACTT 1244
Db      661 CTGGGGCCGTGATGTGCTGTGAAGGACAAGCCAGGCGCTTACACGAGAGTTCACACTT 720
QY      1245 CTTACCTGTGATCGCAGTCACACCAAGAAAGAAATGCGCTGCGCCCTCTGA 1296
Db      721 CTTACCTGTGATCGCAGTCACACCAAGAAAGAAATGCGCTGCGCCCTCTGA 771
```

Search completed: March 21, 2004, 20:50:03
Job time : 3808 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 178.247 Seconds
(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-2

Sequence: 1 MSALLARLLLCVLLVSDSKG.....VSHFLPWRSHRENGAL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1808:*
- 2: geneseqp1908:*
- 3: geneseqp2008:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2394	100.0	431	1	AAp50114
2	2394	100.0	431	1	AAp60783
3	2394	100.0	431	1	AAp70258
4	2394	100.0	431	1	AAp80430
5	2394	100.0	431	1	AAp81204
6	2394	100.0	431	1	AAp92119
7	2394	100.0	431	2	AAp07112
8	2394	100.0	431	2	AAp04253
9	2394	100.0	431	2	AAp63141
10	2394	100.0	431	2	AAp47903
11	2394	100.0	431	3	AAp50869
12	2394	100.0	431	3	AAp95951
13	2394	100.0	431	4	AAp84605
14	2394	100.0	431	5	AAp17128
15	2394	100.0	431	5	AAp99228
16	2394	100.0	431	5	AAp79460
17	2394	100.0	431	5	AAp55855
18	2394	100.0	431	6	AAp56547
19	2394	100.0	431	6	AAp56708
20	2394	100.0	431	6	AAp11076
21	2394	100.0	431	6	AAp892137
22	2394	100.0	434	2	AAp20537
23	2394	100.0	434	2	AAp20538
24	2394	100.0	436	2	AAp20536
25	2392	99.9	431	7	AAp25745

26	2391	99.9	431	1	AAp91886	AAp91886 Sequence
27	2391	99.9	431	1	AAp94764	AAp94764 Non-glyco
28	2391	99.9	431	6	AAp37128	AAp37128 Human uro
29	2391	99.9	431	7	AAp46429	AAp46429 Human pro
30	2390	99.8	431	5	AAp99230	AAp99230 Human pla
31	2389	99.8	431	5	AAp99229	AAp99229 Human pla
32	2386	99.7	431	5	AAp99236	AAp99236 Human pla
33	2385	99.6	431	1	AAp30041	AAp30041 Sequence
34	2384	99.6	431	5	AAp99231	AAp99231 Human pla
35	2383	99.6	431	5	AAp99232	AAp99232 Human pla
36	2382	99.5	431	1	AAp70250	AAp70250 Sequence
37	2382	99.5	431	5	AAp99238	AAp99238 Human pla
38	2381	99.5	431	5	AAp99237	AAp99237 Human pla
39	2380	99.4	431	1	AAp60674	AAp60674 Modified
40	2380	99.4	431	5	AAp99234	AAp99234 Human pla
41	2379	99.4	431	5	AAp99233	AAp99233 Human pla
42	2378	99.3	431	2	AAp33198	AAp33198 Human uri
43	2377	99.3	431	2	AAp33199	AAp33199 Human uri
44	2377	99.3	431	5	AAp99239	AAp99239 Human uri
45	2376	99.2	431	2	AAp33121	AAp33121 Human uri

ALIGNMENTS

RESULT 1	AAp50114	standard; protein; 431 AA.
XX	AAp50114:	
AC	27-SEP-1991 (first entry)	
DT	Sequence encoded by the signal sequence and noncoding region of the pro-	
XX	UK structural gene (Sequence II).	
DE	Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase.	
XX	Homo sapiens.	
OS		
XX		
PH	Key	Location/Qualifiers
FT	Peptide	1..20
FT	Domain	/label= signal peptide
FT	Domain	21..177
FT	Domain	/label= A chain
FT	Domain	179..431
FT	Domain	/label= B chain
XX		
PN	EP154272-A.	
XX	11-SEP-1985.	
PD	23-FEB-1985;	85EP-00102031.
XX	27-FEB-1984;	84JP-00037119.
XX	31-JAN-1985;	85JP-00017969.
PR	(GRC) GREEN CROSS CORP.	
PA	Hiramatsu R, Kaneda T, Nagai M, Rimura H, Nishida M, Suyama T;	
PI	WPI; 1985-224693/37.	
DR	N-PSDB; AAN50138.	
XX		
PT	Glycosylated single-chain pro-urokinase - prepd. by cultivating animal	
XX	cells transfected by DNA prepd. from m RNA.	
PS	Disclosure; Page 8-10; 64pp; English.	
CC	The inventors claim a method of producing single-chain pro-urokinase by	
CC	using as template, mRNA obt. from cells of an established human kidney-	
CC	derived cell line. The urokinase is used to treat thrombosis and embolic	
CC	diseases as well as in the treatment of diseases in combination with	

CC anticancer agents
 XX Sequence 431 AA;
 SQ

Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVLYVSDSKGSELHQPNSNCDLNGCTCVSNKYFNSNIHWNCCKPKKGGQ 60
 DB 1 MRALLARLLLCVLYVSDSKGSELHQPNSNCDLNGCTCVSNKYFNSNIHWNCCKPKKGGQ 60
 QY 61 HCEIDSKTCEYEGNGHFRGKASTDTMGRPCLPMSATVLOOTYHAHRSALQLGLGKN 120
 DB 61 HCEIDSKTCEYEGNGHFRGKASTDTMGRPCLPMSATVLOOTYHAHRSALQLGLGKN 120
 QY 121 YCRNPDNRRPWCYVGVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
 DB 121 YCRNPDNRRPWCYVGVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
 QY 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
 DB 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
 QY 241 RSRINSTOGEMKTEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
 DB 241 RSRINSTOGEMKTEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
 QY 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 DB 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 QY 361 CAADPQWKTDSCQDGGPLVCSLOGRMVLTGIIVSGRCALKDKPGVYTRVSHFLPMIR 420
 DB 361 CAADPQWKTDSCQDGGPLVCSLOGRMVLTGIIVSGRCALKDKPGVYTRVSHFLPMIR 420
 QY 421 SHTKEENGIAL 431
 DB 421 SHTKEENGIAL 431

RESULT 2
 AAP60783
 ID AAP60783 standard; protein; 431 AA.
 XX
 AC AAP60783;
 DT 25-MAR-2003 (revised)
 DT 23-OCT-1991 (first entry)
 XX
 DE Human urokinase.
 KM E.coli; high molecular urokinase.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 21..431
 XX
 XX JP61181377-A.
 XX
 XX 14-AUG-1986.
 XX
 XX 25-JAN-1985; 85JP-00011032.
 XX
 XX 25-JAN-1985; 85JP-00011032.
 XX
 XX (NISC) NISSAN CHEM IND LTD.
 XX (HODO) HODOGAYA CHEM IND CO LTD.
 XX (SAGA) SAGAMI CHEM RES CENTRE.
 XX (CENG) CENTRAL GLASS CO LTD.
 XX (NIPS) NIPPON SODA CO.
 XX (TOYU) TOYO SODA MFG CO LTD.

XX
 DR WPI: 1986-254744/39.
 DR N-PSDB; AAP60703.
 XX
 XX Human urokinase gene - has N-end of aminoacid sequence coded by codon
 PT used in Escherichia coli.
 XX
 PS Disclosure; Fig 2; 19pp; Japanese.
 XX
 CC The claimed gene product may be expressed in a transformed E.coli host,
 CC for the efficient production of high molecular human urokinase. The N-
 CC terminal of the protein expressed by the transforming plasmid is replaced
 CC with a codon frequently used in E.coli. (Updated on 25-MAR-2003 to
 CC correct PA field.)
 CC
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVLYVSDSKGSELHQPNSNCDLNGCTCVSNKYFNSNIHWNCCKPKKGGQ 60
 DB 1 MRALLARLLLCVLYVSDSKGSELHQPNSNCDLNGCTCVSNKYFNSNIHWNCCKPKKGGQ 60
 QY 61 HCEIDSKTCEYEGNGHFRGKASTDTMGRPCLPMSATVLOOTYHAHRSALQLGLGKN 120
 DB 61 HCEIDSKTCEYEGNGHFRGKASTDTMGRPCLPMSATVLOOTYHAHRSALQLGLGKN 120
 QY 121 YCRNPDNRRPWCYVGVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
 DB 121 YCRNPDNRRPWCYVGVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
 QY 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
 DB 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
 QY 241 RSRINSTOGEMKTEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
 DB 241 RSRINSTOGEMKTEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
 QY 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 DB 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 QY 361 CAADPQWKTDSCQDGGPLVCSLOGRMVLTGIIVSGRCALKDKPGVYTRVSHFLPMIR 420
 DB 361 CAADPQWKTDSCQDGGPLVCSLOGRMVLTGIIVSGRCALKDKPGVYTRVSHFLPMIR 420
 QY 421 SHTKEENGIAL 431
 DB 421 SHTKEENGIAL 431

RESULT 3
 AAP70258
 ID AAP70258 standard; protein; 431 AA.
 XX
 AC AAP70258;
 DT 25-MAR-2003 (revised)
 DT 19-MAY-1991 (first entry)
 XX
 DE Sequence of human prourokinase and leader.
 KM Cardiovascular disease treatment; fibrin affinity; thrombolytic; enzyme;
 XX protease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label= leader

FT Protein 21.431
 XX /label= prourokinase
 XX EP231883-A.
 XX 12-AUG-1987.
 PD 29-JAN-1987; 87EP-00101209.
 XX 31-JAN-1986; 86JP-00017734.
 PR 30-JAN-1987; 87JP-00018626.
 XX (SAGA) SAGAMI CHEM RES CENTRE.
 PA (NIPS) NIPPON SODA CO.
 PA (CENG) CENTRAL GLASS CO LTD.
 PA (TOYJ) TOYO SODA MFG CO LTD.
 PA (NISC) NISSAN CHEM IND LTD.
 PA (NISC) NISSAN CHEMICAL INDS KK.
 XX Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;
 PI WPI; 1987-232882/32.
 XX N-PSDB; AAN70390.
 DR Hybrid plasminogen activator-like polypeptide - having a region for
 PT affinity to fibrin from tissue plasminogen activator and a region from
 PT prourokinase.
 XX Disclosure; Fig 2(1-5); 64pp; English.
 PS The TPA portion of the claimed hybrid polypeptide (see FT) may consist of
 CC 2 kringles from N-terminal first serine to 219th glycine of human TPA, 1
 CC kringles from 188th serine to 219th glycine of human TPA or half a kringles
 CC from 161st methionine to 219th glycine (see AAP70257). The C-terminal
 CC half of the hybrid polypeptide may contain an AA sequence from 150th
 CC glutamine to C-terminal 411th leucine of prourokinase (see AAP70258).
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX
 SQ Sequence 431 AA:
 Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALLRLILCTVYVSDSGKSNELHQPNSDCDCLNGSTCVSNKYFSNFIHMCNPKKGGQ 60
 DB 1 MRALLRLILCTVYVSDSGKSNELHQPNSDCDCLNGSTCVSNKYFSNFIHMCNPKKGGQ 60
 QY HCEIDSKTCYEGNGHFYRKGASTDTWGRPCLPMNSATVLAQOTYHAHRSDALQGLGKN 120
 DB HCEIDSKTCYEGNGHFYRKGASTDTWGRPCLPMNSATVLAQOTYHAHRSDALQGLGKN 120
 QY 121 YCNPDPNRRRPMCYVYGLKPLVQECVNDHCDGKKSPPEELKFCQGGKTLRPRKII 180
 DB 121 YCNPDPNRRRPMCYVYGLKPLVQECVNDHCDGKKSPPEELKFCQGGKTLRPRKII 180
 QY 181 GSEFTIENQPMFAIYRRHRGSSVTYVCGSLISPCWVLSATHCFIDYRKEDYIYLG 240
 DB 181 GSEFTIENQPMFAIYRRHRGSSVTYVCGSLISPCWVLSATHCFIDYRKEDYIYLG 240
 QY 241 RSLRLNTOGEMKFEVENLILHSDYADTLAHNDIALALKIRSEKGCAPSRITQITCL 300
 DB 241 RSLRLNTOGEMKFEVENLILHSDYADTLAHNDIALALKIRSEKGCAPSRITQITCL 300
 QY 301 PSMTYNDPQFSTCEITGFGKENSTDYLYPEQLKTTVYKLSHRBCQPHYYGSEVTTKML 360
 DB 301 PSMTYNDPQFSTCEITGFGKENSTDYLYPEQLKTTVYKLSHRBCQPHYYGSEVTTKML 360
 QY 361 CAADPQWKTDSCGDSGGLVCSLOGRMVLTGIVSMRGALCDKDPVYRVSHPLPWIR 420
 DB 361 CAADPQWKTDSCGDSGGLVCSLOGRMVLTGIVSMRGALCDKDPVYRVSHPLPWIR 420
 QY 421 SHTKEENGIAL 431

Db 421 SHTKEENGIAL 431
 RESULT 4
 AAP80430
 ID AAP80430 standard; protein; 431 AA.
 XX
 AC AAP80430;
 XX
 DT 25-MAR-2003 (revised)
 DT 14-SEP-1990 (first entry)
 XX
 DE Deduced AA sequence of the single chain urokinase plasminogen activator
 DE (SCU-PA) cDNA insert prepared from human Hep3 cells.
 XX
 KM Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;
 KM glycerolaldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;
 XX thrombosis treatment.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT Protein 21..411
 XX
 PN EP288435-A.
 PD 26-OCT-1988.
 XX
 PF 11-APR-1988; 88EP-00810234.
 XX
 PR 15-APR-1987; 87GB-00009081.
 PR 16-JUN-1987; 87GB-00014059.
 PR 04-DEC-1987; 87JE-00003299.
 XX
 PA (CIBA) CIBA GEIGY AG.
 XX
 PI Meyhack B, Helm J, Burgi R;
 XX WPI; 1988-301440/43.
 DR N-PSDB; AAN80981.
 XX
 PT Prod. of human single chain urokinase-type plasminogen activator - by
 PT culturing yeast strain transformed with hybrid vector conig. yeast
 PT expression control sequences.
 XX
 PS Example 1; Fig 2; 48pp; English.
 XX
 CC The patent is for the prodn. of human single chain urokinase-type
 CC plasminogen activator (uPA). Mutants of scu-PA are especially those
 CC which render the protein protease resistant. Such scu-PA mutants are
 CC covalently modified at sites of proteolysis by proteases occurring in
 CC blood such as thrombin or plasmin, so that they are no longer susceptible
 CC to protease hydrolysis at these locations. The target sites include
 CC Lys135 to Lys136 (cleavage at this site generates the so-called low
 CC molecular weight form of scu-PA or LUK), Arg156 to Phe157 (susceptible to
 CC thrombin attack) and Lys158 to Ile159 (cleavage at this site by plasmin
 CC generates tuc-PA). Suitable scu-PA mutants have site specific
 CC substitutions, insertions or deletions of residues at one or more of
 CC these target sites. Especially preferred are those mutants in which one
 CC amino acid residue or both amino acid residues forming the target sites
 CC are deleted or in which at least one of these amino acid residues is
 CC replaced by another amino acid residue so that the resulting mutants are
 CC resistant to proteolytic attack. The uPA proteins exhibit the biological
 CC activity of natural human uPA without any refolding procedure being
 CC necessary. They can be used as for known PAs in humans for the prevention
 CC or treatment of thrombosis or other conditions where it is desired to
 CC produce local fibrinolytic or proteolytic activity. (Updated on 25-MAR-
 CC 2003 to correct PR field.)
 XX
 SQ Sequence 431 AA:

Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVTVSDSGSNELHQPNSNCDCLNGCTCVSNKYPSNIHMCNCPKKEGGQ 60
DB 1 MRALLARLLLCVTVSDSGSNELHQPNSNCDCLNGCTCVSNKYPSNIHMCNCPKKEGGQ 60
QY HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
DB HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
QY 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHGGSTVYCGSLISPCWVIAATCFIDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHGGSTVYCGSLISPCWVIAATCFIDYPKKEDYIYVLG 240
QY 241 RSRINSNTQGMKEFEVENTLHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIOTICL 300
DB 241 RSRINSNTQGMKEFEVENTLHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
DB 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCQDSSGGLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSCQDSSGGLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKENGLAL 431
DB 421 SHTKENGLAL 431

```

RESULT 5

AAP81204
 ID AAP81204 standard; protein; 431 AA.

AC AAP81204;

DT 25-MAR-2003 (revised)
 DT 03-DEC-1990 (first entry)

XX Pro-urokinase with signal sequence.

DE pro-urokinase (pro-UK); plasminogen activator; PUK33; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Peptide 1..20 /label= signal peptide

FT Protein 21..431 /label= pro-urokinase

XX EP265874-A.

XX 04-MAY-1988.

XX 23-OCT-1987; 87EP-00115600.

XX 23-OCT-1986; 86JP-00253078.

XX (GREC) GREEN CROSS CORP.

XX Anatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;

XX WPI; 1988-121000/18.

XX N-PSDB; AAN81558.

XX

PT glycosylated single-chain pro-urokinase prodn. - by cultivating DHFR gene
 PT -deficient CHO-K1 cells transformed with a plasmid contg. CDNA, SV40
 promoter and DHFR gene.

XX Disclosure; Page 7; 19p; English.

XX The Arg at position 2 is encoded by TGA(sic). Possible error in the
 CC specification. Should read CGA ? The pro-UK gene was derived from plasmid
 CC PUK33. The cDNA was synthesised using urokinase mRNA isolated from a
 CC human kidney cell line. Pro-UK was cloned into a SV40 promoter-contg.
 CC plasmid, down -stream of the promoter. This plasmid was then ligated to
 CC a DHFR- gene contg. plasmid so that pro-UK and DHFR are inserted in
 CC opposite directions. The recombinant plasmid was used to transform CHO-K1
 CC cell derived DHFR gene-deficient host cells to produce glycosylated
 CC single-chain pro-UK. (Updated on 25-MAR-2003 to correct pf field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVTVSDSGSNELHQPNSNCDCLNGCTCVSNKYPSNIHMCNCPKKEGGQ 60
DB 1 MRALLARLLLCVTVSDSGSNELHQPNSNCDCLNGCTCVSNKYPSNIHMCNCPKKEGGQ 60
QY HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
DB HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQLGLGKN 120
QY 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHGGSTVYCGSLISPCWVIAATCFIDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHGGSTVYCGSLISPCWVIAATCFIDYPKKEDYIYVLG 240
QY 241 RSRINSNTQGMKEFEVENTLHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIOTICL 300
DB 241 RSRINSNTQGMKEFEVENTLHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
DB 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCQDSSGGLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSCQDSSGGLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKENGLAL 431
DB 421 SHTKENGLAL 431

```

RESULT 6

AAP92119
 ID AAP92119 standard; protein; 431 AA.

XX AAP92119;

XX 25-MAR-2003 (revised)

XX 29-JUN-1990 (first entry)

XX Natural human prourokinase.

XX Human prourokinase; antithrombotic; derivative.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "Optional in new deriv."

```
FT Misc-difference 2.155
FT /note= "Incorporated into new deriv."
FT Misc-difference 135
FT /note= "May be replaced by a non-basic AA in new deriv."
FT Misc-difference 156
FT /note= "Undefined residue in new deriv."
FT Misc-difference 157
FT /note= "Pro, Gly, Ala or Val in new deriv."
FT Misc-difference 158
FT /note= "Lys or Arg in new deriv."
FT Misc-difference 158
FT /note= "Lys or Arg in new deriv."
PN MO8901513-A.
XX
XX 23-FEB-1989.
XX
XX 18-AUG-1988; 88MO-JP000815.
XX
XX 19-AUG-1987; 87JP-00204149.
XX
XX (SAGA ) SAGAMI CHEM RES CENTRE.
XX (CENG ) CENTRAL GLASS CO LTD.
XX (HODO ) HODOGAYA CHEM KK.
XX (NIPS ) NIPPON SODA CO.
XX (NISC ) NISSAN CHEM IND LTD.
XX
XX Kobayashi Y, Omori M, Yamada C;
XX PI
XX WPI; 1989-068869/09.
XX DR N-PSDB; AAN91075.
XX
XX Antithrombotic fast-acting pro-urokinase deriv. - produced by culture of
XX PT E.coli transformant contg. new plasmid of PMUT9Q family.
XX PS
XX PS Disclosure; Fig 1; 75pp; Japanese.
XX
XX A human pro-urokinase (PU) deriv. is new which is based upon residues 2-
XX CC 155 of natural human pro-urokinase. The new deriv. is produced by E. coli
XX CC J103/PMUT9Q-RPK in culture. It is a fast-acting drug for the treatment
XX CC and prevention of thrombosis. (Updated on 25-MAR-2003 to correct PA
XX CC field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 2394; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRLLARLLLCVIVSDSKGSLHGVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
DB 1 MRLLARLLLCVIVSDSKGSLHGVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
QY 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
DB 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
QY 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVDCAADGKRPSPPEELKFOCGOKTLRPPFKII 180
DB 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVDCAADGKRPSPPEELKFOCGOKTLRPPFKII 180
QY 181 GGEFTTIENQPMFAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYRKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYRKEDYIVYL 240
QY 241 RSLNLTNTOGEMFVENLILHVDYSDTLAHNDIALKIRSKRGCAPSTTIQITCL 300
DB 241 RSLNLTNTOGEMFVENLILHVDYSDTLAHNDIALKIRSKRGCAPSTTIQITCL 300
QY 301 PSYNDPOFGTSCETIGFGKENSTDYLYPBLKMTVVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSYNDPOFGTSCETIGFGKENSTDYLYPBLKMTVVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCQDGGGPIVCSLQGRMTLTGIVSMRGCCALDKDQGVYTVRVSHFLPMIR 420
DB 361 CAADPQWKTDSCQDGGGPIVCSLQGRMTLTGIVSMRGCCALDKDQGVYTVRVSHFLPMIR 420
```

```
DB 361 CAADPQWKTDSCQDGGGPIVCSLQGRMTLTGIVSMRGCCALDKDQGVYTVRVSHFLPMIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431
RESULT 7
AAR07112
ID AAR07112 standard; protein; 431 AA.
XX
XX AAR07112;
AC
XX 25-MAR-2003 (revised)
DT 24-JAN-1991 (first entry)
XX
XX Human pro-urokinase encoded by plasmid pUK1.
DE
XX pro-urokinase; transgenic mice.
XX
XX Synthetic.
XX OS
XX EP390592-A.
XX
XX 03-OCT-1990.
XX
XX 30-MAR-1990; 90EP-00303445.
XX
XX 31-MAR-1989; 89JP-00078574.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX (EXPE-) CENT INST EXPER ANIMALS.
XX (JIKK-) JIKKEN DOBUTSU CHUO KENK.
XX
XX Sekine S, Ito S, Katsuki M;
XX PI
XX WPI; 1990-299492/40.
XX DR N-PSDB; AAC06049.
XX
XX Prodn. of recombinant protein, esp. human pro-urokinase - from milk of
XX PT transgenic animals using promoter of bovine alpha S1 casein chromosomal
XX PT gene.
XX
XX Example; Table 1; 55pp; English.
XX
XX E.coli strain C600SF8 was transformed with recombinant plasmid containing
XX CC ds DNA derived from human pharynx cancer cell strain Detroit 562. 10000
XX CC colonies were screened and one positive clone was identified. Plasmid
XX CC pUK1 was isolated and found to contain the coding region and 3' non-
XX CC coding region of pro-UK downstream of Cys(41). Four silent substitutions
XX CC were identified (c.f. Holmes et al., Biotechnology, vol.3, p.923 (1985) as
XX CC fellows: (etic) Asn(1254), AAC to AAT; Leu(140), CTA to CTG; Pro(1345), CCC
XX CC to CCA; Gln(1346), CAA to CAG. See also AA006045-006048 and AA006392.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 2394; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRLLARLLLCVIVSDSKGSLHGVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
DB 1 MRLLARLLLCVIVSDSKGSLHGVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
QY 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
DB 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
QY 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVDCAADGKRPSPPEELKFOCGOKTLRPPFKII 180
DB 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVDCAADGKRPSPPEELKFOCGOKTLRPPFKII 180
```


QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGGSLISPCWVISAHCIFIDYPKKEDYIVYLG 240
 DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGGSLISPCWVISAHCIFIDYPKKEDYIVYLG 240
 QY 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 DB 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 QY 301 PSMYNDPQFGTSCETITGFGKENSVDLYPEOLKMTVVLISHRECOQPHYGSSEVTTKML 360
 DB 301 PSMYNDPQFGTSCETITGFGKENSVDLYPEOLKMTVVLISHRECOQPHYGSSEVTTKML 360
 QY 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAIKDKPGVYTRVSHFLPWIR 420
 DB 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAIKDKPGVYTRVSHFLPWIR 420
 QY 421 SHTKEENGIAL 431
 DB 421 SHTKEENGIAL 431

RESULT 8

AA04253
 ID AAR04253 standard; protein; 431 AA.

AC AAR04253;
 XX

DT 25-MAR-2003 (revised)
 DT 12-SEP-1990 (first entry)

XX Human pro-urokinase from the cDNA of clone pCUK176.

XX Non-glycosylated; pro-urokinase; E. coli; Pcrp promoter; MS-2 RBS.

XX Synthetic.

XX EP365894-A.

XX 02-MAY-1990.

XX 06-OCT-1989; 89EP-00118586.

XX 11-OCT-1988; 86GB-00023833.

XX (FARM) FARMITALIA ERBA SPA CARLO.

XX Brandazza A, Sarmientos F, Orsini G;

XX WPI; 1990-133447/18.

XX DR N-PSDB; AAQ04107.

XX Non-glycosylated pro-urokinase prodn. - using E coli B strains and E coli

XX promoter PTrp and Shine-Dalgarno sequence MS-2.

XX Disclosure; Page ?; -pp; English.

XX SRR residue at position 21 is the start of the mature proUK. Non-
 CC glycosylated proUK (MW 45KD) produced by E.coli B strain containing the
 CC sequence. . See also AAQ04101-07. (Updated on 25-MAR-2003 to correct PA
 CC field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLICVAVSDSKGNSNELHOVPSNCDLNGTCVSNKYFSNINHCNCPKKGCG 60
 DB 1 MRALLARLLICVAVSDSKGNSNELHOVPSNCDLNGTCVSNKYFSNINHCNCPKKGCG 60
 QY 61 HCEIDSKTCEBNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQGLGKN 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 61 HCEIDSKTCEBNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQGLGKN 120
 QY 121 YCRNPDNRRRPWCYVQGLPELVQECVHDCADGKSPSPPEELKFOCGQKTLRPRFKII 180
 DB 121 YCRNPDNRRRPWCYVQGLPELVQECVHDCADGKSPSPPEELKFOCGQKTLRPRFKII 180
 QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGGSLISPCWVISAHCIFIDYPKKEDYIVYLG 240
 DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGGSLISPCWVISAHCIFIDYPKKEDYIVYLG 240
 QY 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 DB 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 QY 301 PSMYNDPQFGTSCETITGFGKENSVDLYPEOLKMTVVLISHRECOQPHYGSSEVTTKML 360
 DB 301 PSMYNDPQFGTSCETITGFGKENSVDLYPEOLKMTVVLISHRECOQPHYGSSEVTTKML 360
 QY 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAIKDKPGVYTRVSHFLPWIR 420
 DB 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAIKDKPGVYTRVSHFLPWIR 420
 QY 421 SHTKEENGIAL 431
 DB 421 SHTKEENGIAL 431

RESULT 9

AA063141
 ID AAR63141 standard; protein; 431 AA.

AC AAR63141;
 XX

DT 25-MAR-2003 (revised)
 DT 09-JUN-1995 (first entry)

XX Full length human urokinase protein.

XX Human urokinase glycoproteins; cardiovascular diseases;

XX pulmonary embolism.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Sig peptide 1..20

FT Disulfide-bond 70..151

FT Disulfide-bond 91..133

FT Disulfide-bond 122..146

FT Disulfide-bond 168..299

FT Cleavage-site 179..180
 /note="cleavage of this site produces a bioactive two

FT chain form of urokinase"

FT Disulfide-bond 209..225

FT Disulfide-bond 217..288

FT Disulfide-bond 313..382

FT Disulfide-bond 345..361

FT Disulfide-bond 372..400

PN EP620279-A1.

PD 19-OCT-1994.

PF 14-APR-1983; 94EP-00104777.

PR 15-APR-1982; 82US-00368773.

PR 14-MAR-1983; 83US-00474930.

PR 14-APR-1983; 83EP-00103629.

XX (GETH) GENENTECH INC.

XX Heyneker HL, Holmes WE, Vehar GA;

XX WPI; 1994-318362/40.

DR N-PSDB; AAQ73483.

XX Prod. of human urokinase glycoproteins - using a recombinant expression
XX PT system used for the treatment of vascular diseases or conditions.

XX PS Claim 1, Fig 4, 41pp; English.

XX AAQ73483 is the cDNA sequence which encodes AAR63141 the full length
CC 5400 dalton human urokinase (UK) protein. This cDNA was used in the
CC construction of a plasmid capable of transforming either yeast or
CC vertebrate cells, enabling them to produce the 5400 dalton human UK
CC protein. The UK glycoprotein produced could then be used in the treatment
CC of cardiovascular diseases, including pulmonary embolism. The UK produced
CC using this method had the advantage of a specific activity towards fibrin
CC and extant thrombi, not demonstrated previously with UK isolated from
CC natural sources. (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct
CC PR field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 3.1e-184; Mismatches 0; Indels 0; Gaps 0;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVIVSDSKSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVSDSKSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKN 120
DB 61 HCEIDSKTCYEGNGHFRGKASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKN 120
QY 121 YCRNPDRRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCGQKTLAPRFKII 180
DB 121 YCRNPDRRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCGQKTLAPRFKII 180
QY 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDYRKEDYIYLG 240
DB 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDYRKEDYIYLG 240
QY 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRCAOPSRTIOTICL 300
DB 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRCAOPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETGFGKENSVDYLPQOLKMTVVKLISHRECOQPHYVGSVTTKML 360
DB 301 PSMYNDPOFGTSCETGFGKENSVDYLPQOLKMTVVKLISHRECOQPHYVGSVTTKML 360
QY 361 CAADPOWKTDSCQDGGPLVCSLOGRMTLTGIYSWGRGALKDKPGVYTRVSHFLPMIR 420
DB 361 CAADPOWKTDSCQDGGPLVCSLOGRMTLTGIYSWGRGALKDKPGVYTRVSHFLPMIR 420
QY 421 SHTKEENGALAL 431
DB 421 SHTKEENGALAL 431

RESULT 10

AAAR47903 standard; protein; 431 AA.

AAAR47903;

13-JUL-1994 (first entry)

Pro-urokinase derivative.

Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor.

Homo sapiens.

JP05336965-A.

XX 21-DEC-1993.

XX 17-OCT-1991; 91JP-00269615.

XX 17-OCT-1991; 91JP-00269615.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX WPI; 1994-030907/04.

XX N-PSDB; AAQ55772.

XX Novel human pro-urokinase derivs. having long half-life - with high
XX thrombolytic activity, useful for treatment of thrombosis.

XX Disclosure; Page 15-17; 29pp; Japanese.

XX Sequences (AAQ55771-72) are pro-urokinase derivatives. The products have
XX an inserted sugar moiety having an amino acid substituted, depleted or
XX inserted variant around the thrombin cleavage site. They also have a long
XX half-life allowing them to be used in the treatment of thrombosis

XX Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 3.1e-184; Mismatches 0; Indels 0; Gaps 0;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVIVSDSKSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVSDSKSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKN 120
DB 61 HCEIDSKTCYEGNGHFRGKASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKN 120
QY 121 YCRNPDRRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCGQKTLAPRFKII 180
DB 121 YCRNPDRRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCGQKTLAPRFKII 180
QY 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDYRKEDYIYLG 240
DB 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDYRKEDYIYLG 240
QY 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRCAOPSRTIOTICL 300
DB 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRCAOPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETGFGKENSVDYLPQOLKMTVVKLISHRECOQPHYVGSVTTKML 360
DB 301 PSMYNDPOFGTSCETGFGKENSVDYLPQOLKMTVVKLISHRECOQPHYVGSVTTKML 360
QY 361 CAADPOWKTDSCQDGGPLVCSLOGRMTLTGIYSWGRGALKDKPGVYTRVSHFLPMIR 420
DB 361 CAADPOWKTDSCQDGGPLVCSLOGRMTLTGIYSWGRGALKDKPGVYTRVSHFLPMIR 420
QY 421 SHTKEENGALAL 431
DB 421 SHTKEENGALAL 431

RESULT 11

AAAY50869 standard; protein; 431 AA.

AAAY50869;

24-FEB-2000 (first entry)

Human urokinase protein fragment.

Urokinase; human; thrombolytic agent; streptokinase; antigenic;
blood clot; heart attack; treatment.

```
XX OS Homo sapiens.
XX PN MO9957251-A2.
XX PD 11-NOV-1999.
XX PF 06-MAY-1999; 99WO-US010086.
XX PR 06-MAY-1998; 98US-0084392P.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PI Zhang XC, Lin X, Tang JUN,
XX DR WPI; 2000-052966/04.
XX PT New thrombolytic agents derived from modified humanized streptokinase,
XX PS useful for treating blood clot disorders.
XX PS Disclosure; Page 46-48; 55pp; English.
CC This invention describes a novel thrombolytic agent comprising
CC streptokinase where at least one nonessential portion has been modified.
CC The invention also describes a method of forming a thrombolytic agent
CC comprising determining a nonessential portion of streptokinase and
CC modifying the nonessential portion to render the resulting protein less
CC antigenic. The modified streptokinase is used to treat blood clot
CC disorders, such as heart attacks. The modified streptokinase has less
CC antigenicity than streptokinase but is still able to complex plasminogen
CC and lead to plasminogen activation. Modified streptokinase with the
CC nonessential portions removed or truncated simplify the molecule. Such
CC smaller proteins are cheaper and easier to produce. This sequence
CC represents a fragment of the human urokinase protein which is used in the
CC description of the method of the invention
XX SQ
Sequence 431 AA;
Query Match 100.0%; Score 2394; DB 3; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALLARLLLCVLYVSDSGKSNELHQPNSDCCLNGGTCVSNKFFSNHMCNCPKFFGQ 60
DB 1 MRALLARLLLCVLYVSDSGKSNELHQPNSDCCLNGGTCVSNKFFSNHMCNCPKFFGQ 60
QY 61 HCEIDSKKTCYEGNGHFYRGKASTDTMGPRCLPMNSATVLOQTYHAHRSDALQLGIGKN 120
DB 61 HCEIDSKKTCYEGNGHFYRGKASTDTMGPRCLPMNSATVLOQTYHAHRSDALQLGIGKN 120
QY 121 YCRNPDNRRRPMWCYOVGLKPLVOECWVHDCAHGKPPSPPEELKFCGCGKTLRPRFKII 180
DB 121 YCRNPDNRRRPMWCYOVGLKPLVOECWVHDCAHGKPPSPPEELKFCGCGKTLRPRFKII 180
QY 181 GGEFTTLENQWFAIYRRHRGSSVTYVCCGSLISPCWVISAATCFIDYRKEDYIYVLG 240
DB 181 GGEFTTLENQWFAIYRRHRGSSVTYVCCGSLISPCWVISAATCFIDYRKEDYIYVLG 240
QY 241 RSRINSTQGEEMKEVENLILHKDYSADTLAHNDIALILKIRSEKCAOPSRITQITCL 300
DB 241 RSRINSTQGEEMKEVENLILHKDYSADTLAHNDIALILKIRSEKCAOPSRITQITCL 300
QY 301 PSWYNDQFGTSCITTFGKBNSTDYLYPRLQKMTVVKLISHRECOQPHYYSSEVTTKML 360
DB 301 PSWYNDQFGTSCITTFGKBNSTDYLYPRLQKMTVVKLISHRECOQPHYYSSEVTTKML 360
QY 361 CAADPQKTDSCQSDSGPLVCSLOGRMTLGIYVSWRGALCKDKPVTYRVSHFLPMIR 420
DB 361 CAADPQKTDSCQSDSGPLVCSLOGRMTLGIYVSWRGALCKDKPVTYRVSHFLPMIR 420
QY 421 SHTKEENGAL 431
DB 421 SHTKEENGAL 431
```

```
RESULT 12
ID AA99591
AA99591 standard; protein; 431 AA.
AC AA99591;
DT 13-SEP-2000 (first entry)
DE Human plasminogen activator urokinase, u-PA.
KW Human; serine protease; plasminogen activator; cardiac; thrombolytic;
heart attack; stroke; blood clotting disorder.
OS Homo sapiens.
XX MO200032759-A1.
XX PD 08-JUN-2000.
XX PF 06-MAY-1999; 99WO-US009991.
XX PR 02-DEC-1998; 98US-0110588P.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PI Lin X, Zhang XC, Tang JUN,
XX DR WPI; 2000-422975/36.
XX PT Polypeptide with plasminogen activator activity useful as thrombolytic
XX agent for treating blood clot disorders e.g. heart attack, comprises 10
XX amino acid peptide fragment for recognition or activation of plasminogen.
XX PS Disclosure; Page 26-28; 41pp; English.
XX PS The present sequence is human plasminogen activator urokinase (u-PA), a
XX CC serine protease which hydrolyses a peptide bond in human plasminogen
XX CC to convert it to its active form, plasmin. Plasminogen is the principal
XX CC serine protease zymogen in the extracellular fluids of vertebrates.
XX CC Plasmin is implicated in pericellular proteolysis associated with a wide
XX CC range of physiological and pathological processes. Plasminogen activators
XX CC regulate plasminogen expression either by hydrolysing a peptide bond, as
XX CC in the case of u-PA, or by forming tight binding complexes with
XX CC plasminogen to spontaneously convert it to plasmin. Review of sequence
XX CC homologues of several plasminogen activators and chymotrypsin has
XX CC identified a six amino acid peptide involved in plasminogen activation.
XX CC This peptide is particularly useful when inserted between amino acid
XX CC residues 644 and 645 of full length human plasminogen. Novel plasminogen
XX CC activators have been made based upon the plasminogen
XX CC activation/recognition site of plasminogen binding proteins. The
XX CC polypeptides are useful in preparing thrombolytic agents for treating
XX CC blood clotting disorders such as heart attack
XX SQ
Sequence 431 AA;
Query Match 100.0%; Score 2394; DB 3; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALLARLLLCVLYVSDSGKSNELHQPNSDCCLNGGTCVSNKFFSNHMCNCPKFFGQ 60
DB 1 MRALLARLLLCVLYVSDSGKSNELHQPNSDCCLNGGTCVSNKFFSNHMCNCPKFFGQ 60
QY 61 HCEIDSKKTCYEGNGHFYRGKASTDTMGPRCLPMNSATVLOQTYHAHRSDALQLGIGKN 120
DB 61 HCEIDSKKTCYEGNGHFYRGKASTDTMGPRCLPMNSATVLOQTYHAHRSDALQLGIGKN 120
QY 121 YCRNPDNRRRPMWCYOVGLKPLVOECWVHDCAHGKPPSPPEELKFCGCGKTLRPRFKII 180
DB 121 YCRNPDNRRRPMWCYOVGLKPLVOECWVHDCAHGKPPSPPEELKFCGCGKTLRPRFKII 180
```

QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYRKEDYIYVLG 240
 DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYRKEDYIYVLG 240
 QY 241 RSRINSTQEMKFEVENILIHQYSDTLAHNDIALILKIRSEKGCAPSRITQITCL 300
 DB 241 RSRINSTQEMKFEVENILIHQYSDTLAHNDIALILKIRSEKGCAPSRITQITCL 300
 QY 301 PSMTNDPQFSGTCEITFGKENSTDYLYPQOLKMTVVKLISHRECOQPHYGSEVTTKML 360
 DB 301 PSMTNDPQFSGTCEITFGKENSTDYLYPQOLKMTVVKLISHRECOQPHYGSEVTTKML 360
 QY 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVTVRVSHFLPMIR 420
 DB 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVTVRVSHFLPMIR 420
 QY 421 SHTKEENGIAL 431
 DB 421 SHTKEENGIAL 431

RESULT 13

AAB84605
 ID AAB84605 standard; protein; 431 AA.

AC AAB84605;

DT 05-SEP-2001 (first entry)

DE Amino acid sequence of urokinase plasminogen activator.

XX Growth factor; protein inhibitor; protease; damaged tissue;
 XX platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
 XX connective tissue derived growth factor; CTGF; chrysalin; VEGF;
 XX keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
 XX transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
 XX granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
 XX vascular endothelial growth factor; urokinase plasminogen activator;
 XX dermal ulcer; wound.

OS Homo sapiens.

PN MO200149309-A2.

PD 12-JUL-2001.

PF 21-DEC-2000; 2000MO-IB001935.

PR 29-DEC-1999; 99GB-00030768.

PA (PF12) PFIZER LTD.

PA (PF12) PFIZER INC.

PI Davies MJ, Huggins JP, McIntosh FS, Occlleston NL;

DR WPI; 2001-418351/44.

DR N-PSDB; AAH28220.

PT Composition for the treatment of damaged tissue i.e. chronic wounds and

PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth

PT factor.

PS Disclosure; Page 550; 572pp; English.

XX The specification describes a pharmaceutical composition, comprising a
 XX growth factor, an inhibitor agent, i.e. a protease. The inhibitor agent
 XX inhibits the action of at least one specific adverse protein, i.e. a
 XX protease, that is upregulated in a damaged tissue such as a wound
 XX environment. Growth factors which are included in the composition of the
 XX invention are platelet-derived growth factor (PDGF), fibroblast growth
 XX factor (FGF), connective tissue derived growth factor (CTGF), factor-beta
 XX keratinocyte-derived growth factor (KGF), transforming growth factor (GM-CSF),
 XX (TGF-beta), granulocyte macrophage colony stimulating factor (GM-CSF),

CC epidermal growth factor (EGF), vascular endothelial growth factor (VEGF),
 CC and chrysalin. Inhibitors which are included in the composition of the
 CC invention include inhibitors of urokinase-type plasminogen activator
 CC (uPA) and matrix metalloproteinase (MMP). The composition is useful for
 CC the treatment of chronic damaged tissue, i.e. wounds and dermal ulcers.
 CC The present sequence represents a human uPA, and is used to produce the
 CC composition of the invention

SQ Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 4; Length 431;

Best Local Similarity 100.0%; Pred. No. 3, 1e-184; Indels 0; Gaps 0;

Matches 431; Conservative 0; Mismatches 0;

QY 1 MRALLARLLLCVIVSPDSKSNELHQPNSDCDCLNGTCVSNKYFSNIHMCNCPKKGQ 60

DB 1 MRALLARLLLCVIVSPDSKSNELHQPNSDCDCLNGTCVSNKYFSNIHMCNCPKKGQ 60

QY 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNSATVLOQTYHAHRSALQGLAKEN 120

DB 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNSATVLOQTYHAHRSALQGLAKEN 120

QY 121 YCRNPDRRRPWCYVGLKPLVQECVMDCAQKSSPPEELKPGCGKTLRPRKII 180

DB 121 YCRNPDRRRPWCYVGLKPLVQECVMDCAQKSSPPEELKPGCGKTLRPRKII 180

QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYRKEDYIYVLG 240

DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYRKEDYIYVLG 240

QY 241 RSRINSTQEMKFEVENILIHQYSDTLAHNDIALILKIRSEKGCAPSRITQITCL 300

DB 241 RSRINSTQEMKFEVENILIHQYSDTLAHNDIALILKIRSEKGCAPSRITQITCL 300

QY 301 PSMTNDPQFSGTCEITFGKENSTDYLYPQOLKMTVVKLISHRECOQPHYGSEVTTKML 360

DB 301 PSMTNDPQFSGTCEITFGKENSTDYLYPQOLKMTVVKLISHRECOQPHYGSEVTTKML 360

QY 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVTVRVSHFLPMIR 420

DB 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVTVRVSHFLPMIR 420

QY 421 SHTKEENGIAL 431

DB 421 SHTKEENGIAL 431

RESULT 14

AAE17128 standard; protein; 431 AA.

ID AAE17128

AC AAE17128;

DT 18-APR-2002 (first entry)

DE Human uPA protein.

OS Homo sapiens.

PN MO200196606-A2.

PD 20-DEC-2001.

PF 14-JUN-2001; 2001MO-US019248.

PR 14-JUN-2000; 2000US-00593488.

PA (NYXI-) NYXIS NEURO THERAPIES INC.

XX

PI Yamamoto H, Kroes R, Moskal JR;
XX WPI; 2002-130746/17.
DR N-PSDB; AAD27855.
XX
PT Identifying a compound for treating cancer, comprises detecting
PT transcription factor Ets-1, N-acetylglucosaminyltransferase V, urokinase-
PT type plasminogen activator, matrix-type metalloproteinase-1 and -3 gene
PT expression.
XX
PS Example 1; Page 62-63; 63pp; English.
XX
CC The invention relates to a method of identifying a compound for treating
CC cancer. The method involves detecting the expression of a panel of
CC sequences selected from transcription factor Ets-1, urokinase-type
CC plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),
CC matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method
CC is useful for identifying a compound that affects a cell, particularly a
CC cancer cell or glioma cell, or a cell that is involved in inflammation.
CC It is used for diagnosing and/or treating cancer or other conditions that
CC are affected by one or more members of a panel of genes or their protein
CC product. The method is also useful for drug discovery, drug safety
CC evaluations and in gene therapy. The present sequence is human uPA
CC protein
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 2394; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 3,1e-184; Mismatches 0; Gaps 0;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALLARLLLCVLVVS DSKSGSNELHQPVSNCCLNGTCVSNKYFSNIHMCNCPKFGGQ 60
Db 1 MRALLARLLLCVLVVS DSKSGSNELHQPVSNCCLNGTCVSNKYFSNIHMCNCPKFGGQ 60
QY 61 HCEIDSKTCTCYEGNGHFPYRGKASTDTMGPRCLPWN SATVTLQOQYHAHRS DALQLGLGKRN 120
Db 61 HCEIDSKTCTCYEGNGHFPYRGKASTDTMGPRCLPWN SATVTLQOQYHAHRS DALQLGLGKRN 120
QY 121 YCRNPDNRRRPWCYVOYGLKPLVQECMVHDCADGKKSSPPEELKFCOGGOKTLRPRFKII 180
Db 121 YCRNPDNRRRPWCYVOYGLKPLVQECMVHDCADGKKSSPPEELKFCOGGOKTLRPRFKII 180
QY 181 GGEFTTLENQFWFAIYRRHSGSVTVVCGGSLISPCWVISA TFCFIDYPKKEDYIYVLG 240
Db 181 GGEFTTLENQFWFAIYRRHSGSVTVVCGGSLISPCWVISA TFCFIDYPKKEDYIYVLG 240
QY 241 RSRINSTOGEMKEVENLIIHKYSDTLAHNDIALMLKTRSKRGCAOPSRITQITCL 300
Db 241 RSRINSTOGEMKEVENLIIHKYSDTLAHNDIALMLKTRSKRGCAOPSRITQITCL 300
QY 301 PSMVNDPQFGTSCITGFGKENS TDVLYPEQLKMTVVKLISHRCQOQPHYVGSFVTTKML 360
Db 301 PSMVNDPQFGTSCITGFGKENS TDVLYPEQLKMTVVKLISHRCQOQPHYVGSFVTTKML 360
QY 361 CAADPQWKTDS CGQDSGGLVCSLQGRMTLTGIYSWGRGALKDKPGVYTRVSHFLPWIR 420
Db 361 CAADPQWKTDS CGQDSGGLVCSLQGRMTLTGIYSWGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
Db 421 SHTKEENGIAL 431
RESULT 15
ID AAU99228
AAU99228 standard; protein; 431 AA.
AC AAU99228;
XX 24-SEP-2002 (first entry)
DT Human plasminogen activator, urokinase (PLAU).
XX
DE

XX Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
KM cytostatic; serine protease; thrombolytic disorder; isogene;
KM pulmonary embolism; chromosome 10q24-qter; haplotype; genotype; SNP;
KM single nucleotide polymorphism; thrombolytic; gene therapy.
XX
OS Homo sapiens.
XX
PN W0200240503-A2.
XX
PD 23-MAY-2002.
XX
PF 14-NOV-2001; 2001WO-US044001.
XX
PR 17-NOV-2000; 2000US-0249703P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
PI Anastasio AE, Bentivegna SC, Koshy B;
XX WPI; 2002-519370/55.
DR N-PSDB; ABR6597, ABR6598.
XX
PT Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
PT useful for improving efficiency and reliability in drug development for
PT treating thrombolytic disorders and cancer.
XX
PS Claim 27, Fig 3; 92pp; English.
XX
CC The invention relates to a polynucleotide comprising a first nucleotide
CC sequence (NS1) comprising a PLAU (plasminogen activator, urokinase, a
CC serine protease) isogene selected from isogenes 1-9 and 11-20 given in
CC the specification, where each isogene comprises the regions of the PLAU
CC gene or cDNA and is further defined by the corresponding sequence of
CC polymorphisms (defining single nucleotide polymorphism, SNP). Also
CC included are methods of haplotyping/genotyping (and predicting the
CC haplotype/genotype of the PLAU gene of an individual, identifying an
CC association between a trait and at least one haplotype or haplotype pair
CC of the PLAU gene, an isolated oligonucleotide for detecting a
CC polymorphism in the PLAU gene, a recombinant non-human organism
CC transfected with the gene or cDNA, fragments of the
CC polynucleotides of at least 10 base pairs encompassing a polymorphic
CC site, an isolated polymorphic variant PLAU protein or fragment, an
CC isolated monoclonal antibody specific for PLAU, a computer system for
CC analyzing and analyzing polymorphism data for the PLAU gene and a genome
CC anthology for the PLAU gene. PLAU is useful in screening for drugs
CC targeting PLAU that are useful for treating thrombolytic disorders and
CC cancers. The methods are useful for improving the efficiency and
CC reliability of the discovery and development of drugs for treating
CC diseases associated with PLAU activity, in validating PLAU as a drug
CC target and in the design of clinical trials for treating a specific
CC condition of disease associated with PLAU activity. The antibody is
CC useful in diagnostic, prognostic and therapeutic methods. PLAU
CC polynucleotides are useful in studying the expression and function of
CC PLAU, and in expressing PLAU protein for use in screening for candidate
CC drugs to treat diseases related to PLAU activity. The gene for PLAU is
CC located on chromosome 10q24-qter. The present sequence represents the
CC PLAU protein
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 2394; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 3,1e-184; Mismatches 0; Gaps 0;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALLARLLLCVLVVS DSKSGSNELHQPVSNCCLNGTCVSNKYFSNIHMCNCPKFGGQ 60
Db 1 MRALLARLLLCVLVVS DSKSGSNELHQPVSNCCLNGTCVSNKYFSNIHMCNCPKFGGQ 60
QY 61 HCEIDSKTCTCYEGNGHFPYRGKASTDTMGPRCLPWN SATVTLQOQYHAHRS DALQLGLGKRN 120
Db 61 HCEIDSKTCTCYEGNGHFPYRGKASTDTMGPRCLPWN SATVTLQOQYHAHRS DALQLGLGKRN 120

Qy	121	YCRNPDRRRPWCYVQVGLKPLVQECMVDKADGKSSPPEELKFOCGQKTLRPRFKII	180
Db	121	YCRNPDRRRPWCYVQVGLKPLVQECMVDKADGKSSPPEELKFOCGQKTLRPRFKII	180
Qy	181	GGEFTTENQPMFAIYRRHRGGSVTYVCGSLISPCWVISATHCFIDYPKEDYIYLG	240
Db	181	GGEFTTENQPMFAIYRRHRGGSVTYVCGSLISPCWVISATHCFIDYPKEDYIYLG	240
Qy	241	RSRLNSNTQEMKFEVENLILHKDYSADTLAHHNDIALKIRSKGRCAQPSRTIQTICL	300
Db	241	RSRLNSNTQEMKFEVENLILHKDYSADTLAHHNDIALKIRSKGRCAQPSRTIQTICL	300
Qy	301	PSMYNDPQFGTSCETITGFGKENSSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML	360
Db	301	PSMYNDPQFGTSCETITGFGKENSSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML	360
Qy	361	CAADPQWKTDSCQDSSGGLVCSLQGRMTLGTIVSMRGCALKDKPGVYTRVSHFLPMIR	420
Db	361	CAADPQWKTDSCQDSSGGLVCSLQGRMTLGTIVSMRGCALKDKPGVYTRVSHFLPMIR	420
Qy	421	SHTKEENGLAL 431	
Db	421	SHTKEENGLAL 431	

Search completed: March 18, 2004, 11:16:19
 Job time : 180.247 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:13:05 ; Search time 49.9847 seconds
(without alignments)
445.152 Million cell updates/sec

Title: US-10-076-421-2
Perfect score: 2394
Sequence: 1 MRALLRLLCVLVSDSKG.....VSHFLPWIRSHRENGAL 431

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2394	100.0	431	US-09-101-272G-1	Sequence 1, Appli
2	2394	100.0	431	US-09-101-272G-1	Patent No. 5188829
3	2378.5	99.4	430	5219569-2	Patent No. 5219569
4	2312.5	96.6	430	US-07-942-157A-3	Sequence 3, Appli
5	2301	96.1	411	US-08-087-163-1	Sequence 1, Appli
6	2301	96.1	411	US-08-286-748B-18	Sequence 18, Appli
7	2301	96.1	411	US-08-153-799-18	Sequence 18, Appli
8	2298	96.0	411	US-09-403-736-2	Sequence 2, Appli
9	2291	95.7	411	US-09-181-816-1	Sequence 1, Appli
10	2277	95.1	411	US-08-560-098A-48	Sequence 48, Appli
11	2027.5	84.7	432	US-08-560-098A-47	Sequence 47, Appli
12	2022	84.5	365	US-08-093-741-83	Sequence 83, Appli
13	2022	84.5	365	US-08-720-012-83	Sequence 83, Appli
14	2022	84.5	363	US-08-560-098A-44	Sequence 44, Appli
15	2022	84.5	363	US-08-967-024C-24	Sequence 24, Appli
16	2022	84.5	363	US-08-967-024C-25	Sequence 25, Appli
17	1507.5	63.0	306	US-08-560-098A-45	Sequence 45, Appli
18	1507.5	63.0	331	US-08-560-098A-46	Sequence 46, Appli
19	1382	57.7	253	US-08-944-483-73	Sequence 73, Appli
20	1374	57.4	254	US-08-560-098A-49	Sequence 49, Appli
21	1134	47.4	200	US-09-101-272G-73	Sequence 73, Appli
22	883	36.9	157	US-08-142-590B-25	Sequence 25, Appli
23	875.5	36.3	527	US-08-811-949-39	Sequence 39, Appli
24	868.5	36.3	527	US-07-609-510B-16	Sequence 16, Appli
25	868.5	36.3	527	PCT-US91-01025A-2	Sequence 2, Appli
26	868.5	36.3	527	US-85259-8	Patent No. 5185259
27	868.5	36.3	562	US-08-811-949-43	Sequence 43, Appli

28	868.5	36.3	562	2	US-08-560-098A-50	Sequence 50, Appli
29	868.5	36.3	562	2	US-08-883-795A-38	Sequence 38, Appli
30	868.5	36.3	562	4	US-09-703-695A-4	Sequence 4, Appli
31	868.5	36.3	562	6	5185259-3	Patent No. 5185259
32	868.5	36.3	562	6	5200340-2	Patent No. 5200340
33	868.5	36.3	562	6	5344773-2	Patent No. 5344773
34	867.5	36.2	477	2	US-08-560-098A-51	Sequence 51, Appli
35	867.5	36.2	527	6	5520913-1	Patent No. 5520913
36	861.5	36.0	562	6	5244676-5	Patent No. 5244676
37	857	35.8	208	4	US-09-101-272G-98	Sequence 98, Appli
38	804	33.6	355	2	US-08-811-949-59	Sequence 59, Appli
39	799	33.4	437	2	US-08-811-949-51	Sequence 51, Appli
40	797.5	33.3	472	2	US-08-811-949-63	Sequence 63, Appli
41	796	33.2	437	2	US-08-811-949-57	Sequence 57, Appli
42	793	33.1	138	2	US-08-797-689-12	Sequence 12, Appli
43	793	33.1	138	4	US-09-984-186-12	Sequence 12, Appli
44	789	33.0	437	2	US-08-811-949-55	Sequence 55, Appli
45	788	32.9	194	4	US-09-101-272G-80	Sequence 80, Appli

ALIGNMENTS

RESULT 1									
US-09-101-272G-1									
Sequence 1, Application US/09101272G									
Patent No. 6509445									
GENERAL INFORMATION:									
APPLICANT: Nisgin Food Products Co., Ltd.									
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR									
FILE REFERENCE: Q50979									
CURRENT APPLICATION NUMBER: US/09/101,272G									
CURRENT FILING DATE: 1998-07-08									
PRIOR APPLICATION NUMBER: JP 1059/1996									
PRIOR FILING DATE: 1996-01-08									
NUMBER OF SEQ ID NOS: 107									
SOFTWARE: Patentin version 3.1									
SEQ ID NO 1									
LENGTH: 431									
TYPE: PRT									
ORGANISM: Homo sapiens									
FEATURE:									
NAME/KEY: mat.peptide									
LOCATION: (21)..()									
OTHER INFORMATION:									
NAME/KEY: misc.feature									
LOCATION: (20)..()									
OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)									
US-09-101-272G-1									
Query Match									
Best Local Similarity 100.0%; Score 2394; DB 4; Length 431;									
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MRALLRLLCVLVSDSKGSELHQVPSNCDLNGTCVSNKYFSNINHCNPKKGGQ	60						
DB	1	MRALLRLLCVLVSDSKGSELHQVPSNCDLNGTCVSNKYFSNINHCNPKKGGQ	60						
QY	61	HCEIDSKTCYENGNHPTKASTDTMGRCPLFMSATTVLOQTYHARSALDGLGKHN	120						
DB	61	HCEIDSKTCYENGNHPTKASTDTMGRCPLFMSATTVLOQTYHARSALDGLGKHN	120						
QY	121	YCNPNRRRPPWCYVVGKPLVQECVHDCADGKKSPPEELKFCQCGKTLRPRKII	180						
DB	121	YCNPNRRRPPWCYVVGKPLVQECVHDCADGKKSPPEELKFCQCGKTLRPRKII	180						
QY	181	GGEFTTIENQPMFAATYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYVYL	240						
DB	181	GGEFTTIENQPMFAATYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYVYL	240						
QY	241	RSLNNTQCEMKEFEVNLILHVDYADTLAHNDIALKIRKGRCAQPSRTIQTCL	300						
DB	241	RSLNNTQCEMKEFEVNLILHVDYADTLAHNDIALKIRKGRCAQPSRTIQTCL	300						

```

QY 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLISHRECOQPHYGSEVTTTCL 360
DB 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLISHRECOQPHYGSEVTTTCL 360
QY 361 CAADPOWKTDSGQDSGGLVCSLOGRMTLTGIVSWRGALCKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPOWKTDSGQDSGGLVCSLOGRMTLTGIVSWRGALCKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

```

```

RESULT 2
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1986
; SEQ ID NO:1:
; LENGTH: 431
5188829-1

```

```

Query Match      100.0%; Score 2394; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 2,8e-198;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRALLARLLLCVLYVSDSKGSNELHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
DB 1 MRALLARLLLCVLYVSDSKGSNELHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
DB 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
QY 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDCAQKSPPEELKQCGOKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDCAQKSPPEELKQCGOKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 240
QY 241 RSRINNTQGEKMEFEVENILHKDYSADTLAHNDIALIKIRSKGRCAPSRITQITCL 300
DB 241 RSRINNTQGEKMEFEVENILHKDYSADTLAHNDIALIKIRSKGRCAPSRITQITCL 300
QY 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLISHRECOQPHYGSEVTTTCL 360
DB 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLISHRECOQPHYGSEVTTTCL 360
QY 361 CAADPOWKTDSGQDSGGLVCSLOGRMTLTGIVSWRGALCKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPOWKTDSGQDSGGLVCSLOGRMTLTGIVSWRGALCKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

```

```

RESULT 3
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858

```

```

; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO:2:
; LENGTH: 430
5219569-2

```

```

Query Match      99.4%; Score 2378.5; DB 6; Length 430;
Best Local Similarity 99.8%; Pred. No. 6.1e-197;
Matches 430; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

QY 1 MRALLARLLLCVLYVSDSKGSNELHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
DB 1 MRALLARLLLCVLYVSDSKGSNELHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
DB 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
QY 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDCAQKSPPEELKQCGOKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDCAQKSPPEELKQCGOKTLRPRFKII 179
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 239
QY 241 RSRINNTQGEKMEFEVENILHKDYSADTLAHNDIALIKIRSKGRCAPSRITQITCL 300
DB 241 RSRINNTQGEKMEFEVENILHKDYSADTLAHNDIALIKIRSKGRCAPSRITQITCL 299
QY 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLISHRECOQPHYGSEVTTTCL 360
DB 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLISHRECOQPHYGSEVTTTCL 359
QY 361 CAADPOWKTDSGQDSGGLVCSLOGRMTLTGIVSWRGALCKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPOWKTDSGQDSGGLVCSLOGRMTLTGIVSWRGALCKDKPGVYTRVSHFLPWIR 419
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 430

```

```

RESULT 4
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:

```

NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: TS1108Cont.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)815-6508
TELEFAX: (404)815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..19
OTHER INFORMATION: /label= peptide
OTHER INFORMATION: /note= "WAF signal"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 198..203
OTHER INFORMATION: /label= modified
OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 96.6%; Score 2312.5; DB 1; Length 430;
Best Local Similarity 96.5%; Pred. No. 3e-191;
Matches 416; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

QY 1 MRLLARLLICVLVSDSKSNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFGGQ 60
DB 1 MRLCLIS-LVIGLALAEVALASNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFGGQ 59
QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGKHN 120
DB 60 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGKHN 119
QY 121 YCNPNRRRPMCVQVGLKPLVQECVHDCADGKSPPEELKFCGQKTLRPRKII 180
DB 120 YCNPNRRRPMCVQVGLKPLVQECVHDCADGKSPPEELKFCGQKTLRPRKII 179
QY 181 GGEFTTIENQWPAIYRRHGGSVTVVCGSLISPCWISATHCFIDYRKEDYIVYL 240
DB 180 GGEFTTIENQWPAIYRRHGGSVTVVCGSLISPCWISATHCFIDYRKEDYIVYL 239
QY 241 RSLNSNTOGEMFEVENLIHKDYADTLAHNDIALKIRSEGRCAQPSRTIOTICL 300
DB 240 RSLNSNTOGEMFEVENLIHKDYADTLAHNDIALKIRSEGRCAQPSRTIOTICL 299
QY 301 PSMYNDPOFGTSCETGFGKENSTDYLPOLKMTVVKLISHRECOOPHYGSEVTTKML 360
DB 300 PSMYNDPOFGTSCETGFGKENSTDYLPOLKMTVVKLISHRECOOPHYGSEVTTKML 359
QY 361 CAADPOWKTDCGDSGGLVCSLQGMRTLGIIVSWRGALCKDKPGVYTRVSHFLPMIR 420
DB 360 CAADPOWKTDCGDSGGLVCSLQGMRTLGIIVSWRGALCKDKPGVYTRVSHFLPMIR 419
QY 421 SHTKEENGLAL 431
DB 420 SHTKEENGLAL 430

RESULT 5
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692
GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewlich, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04353/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
US-08-087-163-1

Query Match 96.1%; Score 2301; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-190;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SNEHLOVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFGGQCHCEIDSKTCYEGNGHFRG 80
DB 1 SNEHLOVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFGGQCHCEIDSKTCYEGNGHFRG 60
QY 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGKHNVCNPNRRRPMCVQVGLK 140
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGKHNVCNPNRRRPMCVQVGLK 120
QY 141 PLVQECVHDCADGKSPPEELKFCGQKTLRPRKIIIGSEFTTIENQWPAIYRRH 200
DB 121 PLVQECVHDCADGKSPPEELKFCGQKTLRPRKIIIGSEFTTIENQWPAIYRRH 180
QY 201 RGSSTVTVCGSLISPCWISATHCFIDYRKEDYIVYLSRLNSNTOGEMFEVENLI 260
DB 181 RGSSTVTVCGSLISPCWISATHCFIDYRKEDYIVYLSRLNSNTOGEMFEVENLI 240
QY 261 LHKDYADTLAHNDIALKIRSEGRCAQPSRTIOTICLPSMYNDPOFGTSCETGFGK 320
DB 241 LHKDYADTLAHNDIALKIRSEGRCAQPSRTIOTICLPSMYNDPOFGTSCETGFGK 300
QY 321 ENSTDYLPOLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPOWKTDCGDSGGL 380
DB 301 ENSTDYLPOLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPOWKTDCGDSGGL 360
QY 381 VCSLQGMRTLGIIVSWRGALCKDKPGVYTRVSHFLPMIRSHKEENGLAL 431
DB 361 VCSLQGMRTLGIIVSWRGALCKDKPGVYTRVSHFLPMIRSHKEENGLAL 411

RESULT 6
US-08-286-748B-18
Sequence 18, Application US/08286748B
Patent No. 575842
GENERAL INFORMATION:
APPLICANT: Victor Gurewlich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY

TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Rasse
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-18

Query Match 96.1%; Score 2301; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-190;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 SNEHQVPSNCDCLNGGTCTVSNKTFNSINHMNCCKKFGGCHCEIDSKTCYEGNGHFYRG 80
DB 1 SNEHQVPSNCDCLNGGTCTVSNKTFNSINHMNCCKKFGGCHCEIDSKTCYEGNGHFYRG 60
QY 81 KASTDTMGRPCLPWNASATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYOVGLK 140
DB 61 KASTDTMGRPCLPWNASATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYOVGLK 120
QY 141 PLYOECMVHDCADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQWPAIYRRH 200
DB 121 PLYOECMVHDCADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQWPAIYRRH 180
QY 201 RGGSVTVVCGGSLISPCWVISAHCFTIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 260
DB 181 RGGSVTVVCGGSLISPCWVISAHCFTIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 240
QY 261 LHKDYSADTLAHNDIALKIRSKRCAOPSRITQITCLPSWYNDPOFGTSCBITGFGK 320
DB 241 LHKDYSADTLAHNDIALKIRSKRCAOPSRITQITCLPSWYNDPOFGTSCBITGFGK 300
QY 321 ENSTDYLYPEQLKMTVYKLSHRECOQPHYGSEVTTKMLCAADPOWKTDSCGGDSGGLP 380
DB 301 ENSTDYLYPEQLKMTVYKLSHRECOQPHYGSEVTTKMLCAADPOWKTDSCGGDSGGLP 360
QY 381 VCSLOGMTLGTIVSWRGALXDKRQEVYTRVSHFLPWRISHTKEBNGAL 431
DB 361 VCSLOGMTLGTIVSWRGALXDKRQEVYTRVSHFLPWRISHTKEBNGAL 411

RESULT 7
US-08-153-799-18

Sequence 18, Application US/08153799
Patent No. 5766883
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-153-799-18

Query Match 96.1%; Score 2301; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-190;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 SNEHQVPSNCDCLNGGTCTVSNKTFNSINHMNCCKKFGGCHCEIDSKTCYEGNGHFYRG 80
DB 1 SNEHQVPSNCDCLNGGTCTVSNKTFNSINHMNCCKKFGGCHCEIDSKTCYEGNGHFYRG 60
QY 81 KASTDTMGRPCLPWNASATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYOVGLK 140
DB 61 KASTDTMGRPCLPWNASATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYOVGLK 120
QY 141 PLYOECMVHDCADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQWPAIYRRH 200
DB 121 PLYOECMVHDCADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQWPAIYRRH 180
QY 201 RGGSVTVVCGGSLISPCWVISAHCFTIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 260
DB 181 RGGSVTVVCGGSLISPCWVISAHCFTIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 240
QY 261 LHKDYSADTLAHNDIALKIRSKRCAOPSRITQITCLPSWYNDPOFGTSCBITGFGK 320

Db 241 LKHDYADTLAHNDIALKIRSEKRCADPSRTIQTICLPMSYNDPQFSGCEITGFGK 300
 Qy 321 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQKTDSCGDSGGL 380
 Db 301 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQKTDSCGDSGGL 360
 Qy 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 431
 Db 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 411

RESULT 8
 US-09-403-736-2
 ; Sequence 2, Application US/09403736
 ; Patent No. 6638502
 ; GENERAL INFORMATION:

; APPLICANT: Aventis S.A.
 ; APPLICANT: LI, Hong
 ; APPLICANT: LU, He
 ; APPLICANT: GRISCELLI, Frank
 ; APPLICANT: OPOLO, Paule
 ; APPLICANT: SORIA, Claudine
 ; APPLICANT: RAGOT, Thierry
 ; APPLICANT: LEGRAND, Yves
 ; APPLICANT: SORIA, Jeanette
 ; APPLICANT: MABILLAT, Christelle
 ; APPLICANT: PERRICAUDET, Michel
 ; APPLICANT: YEH, Patrice
 ; TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery of An Angiogenesis Anta
 ; FILE REFERENCE: A2778A-US
 ; CURRENT APPLICATION NUMBER: US/09/403, 736
 ; PRIOR FILING DATE: 1999-10-26
 ; PRIOR APPLICATION NUMBER: PCT/EP98/02491
 ; PRIOR FILING DATE: 1998-11-05
 ; PRIOR APPLICATION NUMBER: 60/044, 980
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 411
 ; TYPE: PRT
 ; ORGANISM: humanuokinaae
 ; US-09-403-736-2

Query Match 96.0%; Score 2298; DB 4; Length 411;
 Best Local Similarity 99.8%; Pred. No. 5e-190;
 Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 21 SNEHQPNSNCDCLNGCTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 80
 Db 1 SNEHQPNSNCDCLNGCTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 60
 Qy 81 KASTDTMGRPCLPWNSATVLOQTYHARSALQGLGKNYCRNPNRRPWCYOVGLK 140
 Db 61 KASTDTMGRPCLPWNSATVLOQTYHARSALQGLGKNYCRNPNRRPWCYOVGLK 120
 Qy 141 PLVQECMWDCAADGKRPSPPEELKFCQCGOKTLRPRFKIIGSEFTTIENQPMFAIYRRH 200
 Db 121 PLVQECMWDCAADGKRPSPPEELKFCQCGOKTLRPRFKIIGSEFTTIENQPMFAIYRRH 180
 Qy 201 RGSSTVYVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOGEEMKFEVENLI 260
 Db 181 RGSSTVYVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOGEEMKFEVENLI 240
 Qy 261 LKHDYADTLAHNDIALKIRSEKRCADPSRTIQTICLPMSYNDPQFSGCEITGFGK 320
 Db 241 LKHDYADTLAHNDIALKIRSEKRCADPSRTIQTICLPMSYNDPQFSGCEITGFGK 300
 Qy 321 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQKTDSCGDSGGL 380
 Db 301 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQKTDSCGDSGGL 360

Qy 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 431
 Db 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 411

RESULT 9
 US-09-181-816-1
 ; Sequence 1, Application US/09181816
 ; Patent No. 6277818
 ; GENERAL INFORMATION:

; APPLICANT: MAZAR, Andrew P.
 ; APPLICANT: JONES, Terence R.
 ; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
 ; FILE REFERENCE: 329042000300 SIDN 1-7
 ; CURRENT APPLICATION NUMBER: US/09/181, 816
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 411
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-181-816-1

Query Match 95.7%; Score 2291; DB 3; Length 411;
 Best Local Similarity 99.8%; Pred. No. 2e-189;
 Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 SNEHQPNSNCDCLNGCTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 80
 Db 1 SNEHQPNSNCDCLNGCTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 60
 Qy 81 KASTDTMGRPCLPWNSATVLOQTYHARSALQGLGKNYCRNPNRRPWCYOVGLK 140
 Db 61 KASTDTMGRPCLPWNSATVLOQTYHARSALQGLGKNYCRNPNRRPWCYOVGLK 120
 Qy 141 PLVQECMWDCAADGKRPSPPEELKFCQCGOKTLRPRFKIIGSEFTTIENQPMFAIYRRH 200
 Db 121 PLVQECMWDCAADGKRPSPPEELKFCQCGOKTLRPRFKIIGSEFTTIENQPMFAIYRRH 180
 Qy 201 RGSSTVYVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOGEEMKFEVENLI 260
 Db 181 RGSSTVYVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOGEEMKFEVENLI 240
 Qy 261 LKHDYADTLAHNDIALKIRSEKRCADPSRTIQTICLPMSYNDPQFSGCEITGFGK 320
 Db 241 LKHDYADTLAHNDIALKIRSEKRCADPSRTIQTICLPMSYNDPQFSGCEITGFGK 300
 Qy 321 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQKTDSCGDSGGL 380
 Db 301 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQKTDSCGDSGGL 360
 Qy 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 431
 Db 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 411

RESULT 10
 US-08-560-098A-48
 ; Sequence 48, Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WENDET, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proelins having fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington

STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-48

Query Match 95.1%; Score 2277; DB 2; Length 411;
Best Local Similarity 99.3%; Pred. No. 3.2e-188;
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 21 SNEIHOVPSNDCCLNGGTCVSNKYSNIMHCNCKKFGGHCETDKSKTCYEGNGHFRG 80
DB 1 SNEIHOVPSNDCCLNGGTCVSNKYSNIMHCNCKKFGGHCETDKSKTCYEGNGHFRG 60
QY 81 KASTIDWGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRRPMCVQVGLK 140
DB 61 KASTIDWGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRRPMCVQVGLK 120
QY 141 PLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIEQNPFAIYRRH 200
DB 121 PLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIEQNPFAIYRRH 180
QY 201 RGSQSVTVVCGSLISPCWISATHCFTIDYPKKEDYIVYIGRSRLNSNTQGBMKFEVENLI 260
DB 181 RGSQSVTVVCGSLISPCWISATHCFTIDYPKKEDYIVYIGRSRLNSNTQGBMKFEVENLI 240
QY 261 LHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETGFGK 320
DB 241 LHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETGFGK 300
QY 321 ENSGNDYLYPQQLKMTTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTSCQDSSGGL 380
DB 301 ENSGNDYLYPQQLKMTTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTSCQDSSGGL 360
QY 381 VCSIQGSMTLTGIVSMRGKCALKDKPGVYTRVSHFLPMIRSHTEENGLA 431
DB 361 VCSIQGSMTLTGIVSMRGKCALKDKPGVYTRVSHFLPMIRSHTEENGLA 411

RESULT 11
US-08-560-098A-47
Sequence 47, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDELT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
Coagulation-inhibiting Properties

NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKee, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-47

Query Match 84.7%; Score 2027.5; DB 2; Length 432;
Best Local Similarity 88.8%; Pred. No. 1.1e-166;
Matches 374; Conservative 11; Mismatches 19; Indels 17; Gaps 3;

QY 11 CVLVVSDKSGSNELHGVPSNDCCLNGGTCVSNKYSNIMHCNCKKFGGHCETDKSKTC 70
DB 29 CVLVVSDKSGSNELHGVPSNDCCLNGGTCVSNKYSNIMHCNCKKFGGHCETDKSKTC 71
QY 71 YEGNGHFRGKASTDWMGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRR 130
DB 72 YEGNGHFRGKASTDWMGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRR 131
QY 131 PMCVQVGLKPLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIEQ 190
DB 132 PMCVQVGLKPLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIEQ 191
QY 191 PMFAIYRRHRRGGSVTVVCGSLISPCWISATHCFTIDYPKKEDYIVYIGRSRLNSNTQ 250
DB 192 PMFAIYRRHRRGGSVTVVCGSLISPCWISATHCFTIDYPKKEDYIVYIGRSRLNSNTQ 251
QY 251 EMKFEVENLIHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIQTICLPSMYNDPQFG 310
DB 252 EMKFEVENLIHKDYSADTLAHNDIALKIKRSKGRCAQPSRTIQTICLPSMYNDPQFG 311
QY 311 TSCETGFGKENSSTVLYPQQLKMTTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKT 370
DB 312 TSCETGFGKENSSTVLYPQQLKMTTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKT 371
QY 371 SCQDSSGGLVCSIQGSMTLTGIVSMRGKCALKDKPGVYTRVSHFLPMIRSHTEENGLA 430
DB 372 SCQDSSGGLVCSIQGSMTLTGIVSMRGKCALKDKPGVYTRVSHFLPMIRSHTEENGLA 431
QY 431 L 431
DB 432 L 432

RESULT 12

QY 307 POFSTCEITGFKENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQ 366
DB 241 POFSTCEITGFKENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQ 300
QY 367 WKTSCQSGSGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKEE 426
DB 301 WKTSCQSGSGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKEE 360
QY 427 NGLAL 431
DB 361 NGLAL 365

RESULT 14
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-44

Query Match 84.5%; Score 2022; DB 2; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.9e-166; Indels 0; Gaps 0;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHN YCRND 126
DB 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHN YCRND 61
QY 127 NRRPWCYOVGLKPLVQGCWYHDCADGKRPSSPEELKFOCGGKTLRPRFKIIGGEFTT 186
DB 62 NRRPWCYOVGLKPLVQGCWYHDCADGKRPSSPEELKFOCGGKTLRPRFKIIGGEFTT 121
QY 187 IENQPFPAATYRRHRS GSVTVVCGSLISPCWVISA THCFIDYKKEDEYIVYIGRS LNS 246
DB 122 IENQPFPAATYRRHRS GSVTVVCGSLISPCWVISA THCFIDYKKEDEYIVYIGRS LNS 181

QY 247 NTOGEMKFEVENTILHKVYSA DTLAHNDIALKIRSKERCAQPSRTIOTICLP SWYND 306
DB 182 NTOGEMKFEVENTILHKVYSA DTLAHNDIALKIRSKERCAQPSRTIOTICLP SWYND 241
QY 307 POFSTCEITGFKENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQ 366
DB 242 POFSTCEITGFKENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQ 301
QY 367 WKTSCQSGSGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKEE 426
DB 302 WKTSCQSGSGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKEE 361
QY 427 NGLAL 431
DB 362 NGLAL 366

RESULT 15
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOCNA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-24

Query Match 84.5%; Score 2022; DB 3; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.9e-166; Indels 0; Gaps 0;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHN YCRND 126
DB 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHN YCRND 61
QY 127 NRRPWCYOVGLKPLVQGCWYHDCADGKRPSSPEELKFOCGGKTLRPRFKIIGGEFTT 186

```

Db      62  NRRRPMCTVGVGLKPLVQECMHDCADGKKPSSPPEELKFOCGOKTILPRPKIIGGEFTT 121
Qy      187  IENQPMFAIYRRHRGGSVTYVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRRLNS 246
Db      122  IENQPMFAIYRRHRGGSVTYVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRRLNS 181
Qy      247  NTQEMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIOTICLPSMYND 306
Db      182  NTQEMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIOTICLPSMYND 241
Qy      307  POFGTSCETGPKENSTDYLPBQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQ 366
Db      242  POFGTSCETGPKENSTDYLPBQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQ 301
Qy      367  WKTDSCQSDSGPLVCSLQGRMTLTGIVSNRGCALKDKPGVYTRVSHFLPWIRSHTKEE 426
Db      302  WKTDSCQSDSGPLVCSLQGRMTLTGIVSNRGCALKDKPGVYTRVSHFLPWIRSHTKEE 361
Qy      427  NGIAL 431
Db      362  NGIAL 366

```

Search completed: March 18, 2004, 11:21:21
 Job time : 50.9847 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:19:25 ; Search time 128.263 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-2
Perfect score: 2394
Sequence: 1 MRLRLRLLCVLVSDSKG.....VSHFLPMSHTKENGAL 431

Scoring table: BLOSUM62
Gapco 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895339 residues
Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2394	100.0	431	US-10-411-037-34	Sequence 34, Appl
2	2394	100.0	431	US-10-076-421-2	Sequence 2, Appl
3	2394	100.0	431	US-10-171-311-184	Sequence 184, Appl
4	2394	100.0	431	US-10-301-822-161	Sequence 161, Appl
5	2394	100.0	431	US-10-131-985-21	Sequence 21, Appl
6	2394	100.0	431	US-10-295-027-414	Sequence 414, Appl
7	2394	100.0	431	US-10-295-027-1275	Sequence 1275, Appl
8	2394	99.9	431	US-10-247-671-149	Sequence 149, Appl
9	2391	99.9	431	US-10-193-656-4	Sequence 4, Appl
10	2376	99.2	431	US-09-264-468B-1	Sequence 1, Appl
11	2317	96.8	445	US-10-360-101-266	Sequence 266, Appl
12	2301	96.1	411	US-09-880-503-3	Sequence 3, Appl
13	2291	95.7	411	US-10-407-821-2	Sequence 2, Appl
14	2243	93.7	403	US-09-880-503-6	Sequence 6, Appl
15	1737	72.6	323	US-09-880-503-7	Sequence 7, Appl

16	1628	68.0	337	US-10-106-698-6266	Sequence 6266, Appl
17	1628	68.0	337	US-10-264-049-2927	Sequence 2927, Appl
18	1508	63.0	276	US-09-880-503-5	Sequence 5, Appl
19	1464	61.2	268	US-10-407-821-3	Sequence 3, Appl
20	1333	55.7	246	US-09-264-468B-2	Sequence 2, Appl
21	1318	55.1	241	US-09-898-837A-47	Sequence 47, Appl
22	1096.5	45.8	650	US-10-401-077-1	Sequence 1, Appl
23	868.5	36.3	527	US-09-987-457-18	Sequence 18, Appl
24	868.5	36.3	527	US-09-987-455-19	Sequence 19, Appl
25	868.5	36.3	527	US-10-360-101-203	Sequence 203, Appl
26	868.5	36.3	562	US-09-969-271-7	Sequence 7, Appl
27	868.5	36.3	562	US-09-974-298-145	Sequence 145, Appl
28	868.5	36.3	562	US-10-411-037-26	Sequence 26, Appl
29	868.5	36.3	562	US-10-193-656-8	Sequence 8, Appl
30	868.5	36.3	562	US-10-443-701-4	Sequence 4, Appl
31	837	35.0	143	US-09-880-503-8	Sequence 8, Appl
32	793	33.1	135	US-09-880-503-4	Sequence 4, Appl
33	793	33.1	138	US-09-984-186-12	Sequence 12, Appl
34	793	33.1	138	US-10-237-667-12	Sequence 12, Appl
35	793	33.1	138	US-10-237-708-12	Sequence 12, Appl
36	793	33.1	138	US-10-237-866-12	Sequence 12, Appl
37	793	33.1	138	US-10-237-871-12	Sequence 12, Appl
38	793	33.1	138	US-10-237-624-12	Sequence 12, Appl
39	785.5	32.8	372	US-09-084-491A-3	Sequence 3, Appl
40	785.5	32.8	372	US-10-102-704-3	Sequence 3, Appl
41	780	32.6	354	US-09-987-457-10	Sequence 10, Appl
42	780	32.6	354	US-09-987-455-11	Sequence 11, Appl
43	780	32.6	377	US-09-987-455-8	Sequence 8, Appl
44	741	31.0	343	US-09-987-457-14	Sequence 14, Appl
45	741	31.0	343	US-09-987-455-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-411-037-34
Sequence 34, Application US/10411037
Publication No. US20040043446A1
GENERAL INFORMATION:
APPLICANT: Neose Technologies, Inc.
APPLICANT: Defries, Shawn
APPLICANT: Zopf, David
APPLICANT: Bayer, Robert
APPLICANT: Hakes, David
APPLICANT: Chen, Xi
APPLICANT: Bove, Caryn
TITLE OF INVENTION: ALPHA GALACTOSIDASE A
FILE REFERENCE: 040853-01-5082
CURRENT APPLICATION NUMBER: US/10/411,037
CURRENT FILING DATE: 2003-04-09
PRIOR APPLICATION NUMBER: US 60/328,523
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/344,692
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/387,292
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: US 60/391,777
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/396,594
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: US 60/404,249
PRIOR FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: US 60/407,527
PRIOR FILING DATE: 2002-08-28
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-411-037-34

Query Match 100.0%; Score 2394; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRALLARLLCVLVSDSKSGSNELHQPNSNCDCLNGGTCVSNKXFSNIHWCNCPKFKGGQ 60
DB 1 MRALLARLLCVLVSDSKSGSNELHQPNSNCDCLNGGTCVSNKXFSNIHWCNCPKFKGGQ 60
QY HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
DB HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
QY HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
DB HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
QY YCRNPDNRPRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
DB YCRNPDNRPRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
QY GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKEDYIYVLG 240
DB GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKEDYIYVLG 240
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKEDYIYVLG 240
QY 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
DB 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
DB 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCCGDSGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
DB 361 CAADPQWKTDSCCGDSGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
QY 421 SHTKEENGLAL 431
DB 421 SHTKEENGLAL 431
```

RESULT 2

```
US-10-076-421-2
; Sequence 2, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, MAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2
```

Query Match 100.0%; Score 2394; DB 13; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRALLARLLCVLVSDSKSGSNELHQPNSNCDCLNGGTCVSNKXFSNIHWCNCPKFKGGQ 60
DB 1 MRALLARLLCVLVSDSKSGSNELHQPNSNCDCLNGGTCVSNKXFSNIHWCNCPKFKGGQ 60
QY HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
DB HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
QY 61 HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
DB 61 HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
```

```
QY 121 YCRNPDNRPRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
DB 121 YCRNPDNRPRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKEDYIYVLG 240
QY 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
DB 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
DB 301 PSMYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCCGDSGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
DB 361 CAADPQWKTDSCCGDSGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
QY 421 SHTKEENGLAL 431
DB 421 SHTKEENGLAL 431
```

RESULT 3

```
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US2003008270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatz, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MEI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184
```

Query Match 100.0%; Score 2394; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRALLARLLCVLVSDSKSGSNELHQPNSNCDCLNGGTCVSNKXFSNIHWCNCPKFKGGQ 60
DB 1 MRALLARLLCVLVSDSKSGSNELHQPNSNCDCLNGGTCVSNKXFSNIHWCNCPKFKGGQ 60
QY HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
DB HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
QY 61 HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
DB 61 HCEIDSKTCEBNGHGYRGAASDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
QY 121 YCRNPDNRPRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
DB 121 YCRNPDNRPRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
```

```

QY 181 GGEFTTIENOPWPAIYRRHGGSVTVVCGGSLISPCWVISAHCIFIDYPKKEDYIYLG 240
DB 181 GGEFTTIENOPWPAIYRRHGGSVTVVCGGSLISPCWVISAHCIFIDYPKKEDYIYLG 240
QY 241 RSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALALKIRSKRGCAOPSRTIOTICL 300
DB 241 RSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALALKIRSKRGCAOPSRTIOTICL 300
QY 301 PSMTNDPOFGTSCETIGFGKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSMTNDPOFGTSCETIGFGKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYSWGRGALDKDKPGVTVRVSHPFLPMIR 420
DB 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYSWGRGALDKDKPGVTVRVSHPFLPMIR 420
QY 421 SHTKEENGAL 431
DB 421 SHTKEENGAL 431

```

RESULT 4
US-10-301-822-161

```

; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangl
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEMO1-029P22NM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PaetSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

```

Query Match 100.0%; Score 2394; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKGQ 60
DB 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
QY 121 YCNPNDNRPRPWCYVVGKLPVQECMVHDCADGKSPPEELKFCQCGQKTLRPRFKII 180
DB 121 YCNPNDNRPRPWCYVVGKLPVQECMVHDCADGKSPPEELKFCQCGQKTLRPRFKII 180
QY 181 GGEFTTIENOPWPAIYRRHGGSVTVVCGGSLISPCWVISAHCIFIDYPKKEDYIYLG 240
DB 181 GGEFTTIENOPWPAIYRRHGGSVTVVCGGSLISPCWVISAHCIFIDYPKKEDYIYLG 240

```

```

DB 181 GGEFTTIENOPWPAIYRRHGGSVTVVCGGSLISPCWVISAHCIFIDYPKKEDYIYLG 240
QY 241 RSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALALKIRSKRGCAOPSRTIOTICL 300
DB 241 RSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALALKIRSKRGCAOPSRTIOTICL 300
QY 301 PSMTNDPOFGTSCETIGFGKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSMTNDPOFGTSCETIGFGKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYSWGRGALDKDKPGVTVRVSHPFLPMIR 420
DB 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYSWGRGALDKDKPGVTVRVSHPFLPMIR 420
QY 421 SHTKEENGAL 431
DB 421 SHTKEENGAL 431

```

RESULT 5
US-10-131-985-21

```

; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

```

Query Match 100.0%; Score 2394; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKGQ 60
DB 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
QY 121 YCNPNDNRPRPWCYVVGKLPVQECMVHDCADGKSPPEELKFCQCGQKTLRPRFKII 180
DB 121 YCNPNDNRPRPWCYVVGKLPVQECMVHDCADGKSPPEELKFCQCGQKTLRPRFKII 180
QY 181 GGEFTTIENOPWPAIYRRHGGSVTVVCGGSLISPCWVISAHCIFIDYPKKEDYIYLG 240
DB 181 GGEFTTIENOPWPAIYRRHGGSVTVVCGGSLISPCWVISAHCIFIDYPKKEDYIYLG 240
QY 241 RSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALALKIRSKRGCAOPSRTIOTICL 300
DB 241 RSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALALKIRSKRGCAOPSRTIOTICL 300
QY 301 PSMTNDPOFGTSCETIGFGKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSMTNDPOFGTSCETIGFGKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360

```

QY 361 CAADPQWKTDSGCGDSGGLVCSLQGRMTLTGIVSWGRGALKDKPQVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSGCGDSGGLVCSLQGRMTLTGIVSWGRGALKDKPQVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

RESULT 6
US-10-295-027-414
Sequence 414, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Glsh, Kurt C.
APPLICANT: Glyne, Richard
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: US/10/295,027
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 414
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-414

Query Match 100.0%; Score 2394; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200; Indels 0; Gaps 0;
Matches 431; Conservative 0; Mismatches 0;

QY 1 MRALLARLLLCVLLVSDSKSGNELHQVPSNDCCLNGSTCVSNKXFFSNIHMCNCPKFKGQ 60
DB 1 MRALLARLLLCVLLVSDSKSGNELHQVPSNDCCLNGSTCVSNKXFFSNIHMCNCPKFKGQ 60
QY 61 HCEIDSKTCYEGNGHYRKGASTDTMGRBCLPNSATVYQOTYHARSALQJGJKN 120
DB 61 HCEIDSKTCYEGNGHYRKGASTDTMGRBCLPNSATVYQOTYHARSALQJGJKN 120
QY 121 YCRNPDRRRPWCYVGLKPLVQECMVHDCADGKKSSPEELKFCGQKTLRPRFKII 180
DB 121 YCRNPDRRRPWCYVGLKPLVQECMVHDCADGKKSSPEELKFCGQKTLRPRFKII 180

DB 121 YCRNPDRRRPWCYVGLKPLVQECMVHDCADGKKSSPEELKFCGQKTLRPRFKII 180
QY 181 GGEFTTENDPWFALYRRRGSSVTVYCGSILSPCWYTSATHCFLDPKEDYTVYG 240
DB 181 GGEFTTENDPWFALYRRRGSSVTVYCGSILSPCWYTSATHCFLDPKEDYTVYG 240
QY 241 RSRNSNTQCEMKFEVENLILHROYSADTLAHNDIALKIRSEKRCAPSRFTIOTICL 300
DB 241 RSRNSNTQCEMKFEVENLILHROYSADTLAHNDIALKIRSEKRCAPSRFTIOTICL 300
QY 301 PSMYNDPQGTSCETIFGKENSVDLYPDLKMTYVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSMYNDPQGTSCETIFGKENSVDLYPDLKMTYVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPQWKTDSGCGDSGGLVCSLQGRMTLTGIVSWGRGALKDKPQVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSGCGDSGGLVCSLQGRMTLTGIVSWGRGALKDKPQVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

RESULT 7
US-10-295-027-1275
Sequence 1275, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Glsh, Kurt C.
APPLICANT: Glyne, Richard
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: US/10/295,027
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1275
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-1275

Query Match 100.0%; Score 2394; DB 15; Length 431;

Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRALLARLLLCVIVVSDSKSGNELHQVPSNCDCLNGGTCVSNKXFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVVSDSKSGNELHQVPSNCDCLNGGTCVSNKXFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
QY 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
DB 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
QY 241 RSRINSTQEMKEFEVENLILHVDYADTLAHNDIALALKIRSEKRCAPSRITQITCL 300
DB 241 RSRINSTQEMKEFEVENLILHVDYADTLAHNDIALALKIRSEKRCAPSRITQITCL 300
QY 301 PSYNDPQFGTSCGFTGFGKNSDTLYYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSYNDPQFGTSCGFTGFGKNSDTLYYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
DB 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
QY 421 SHTKEENGALAL 431
DB 421 SHTKEENGALAL 431

RESULT 8
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ. ID NOS: 186
; SOFTWARE: PERL Program
; SEQ. ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149
```

Query Match 99.8%; Score 2392; DB 14; Length 431;
Best Local Similarity 99.8%; Pred. No. 1.6e-200;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRALLARLLLCVIVVSDSKSGNELHQVPSNCDCLNGGTCVSNKXFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVVSDSKSGNELHQVPSNCDCLNGGTCVSNKXFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
QY 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
DB 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
```

```
QY 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
DB 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
QY 241 RSRINSTQEMKEFEVENLILHVDYADTLAHNDIALALKIRSEKRCAPSRITQITCL 300
DB 241 RSRINSTQEMKEFEVENLILHVDYADTLAHNDIALALKIRSEKRCAPSRITQITCL 300
QY 301 PSYNDPQFGTSCGFTGFGKNSDTLYYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSYNDPQFGTSCGFTGFGKNSDTLYYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
DB 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
QY 421 SHTKEENGALAL 431
DB 421 SHTKEENGALAL 431
```

```
RESULT 9
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAL, Rikard
; APPLICANT: Li, Jinnan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ. ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ. ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4
```

Query Match 99.8%; Score 2391; DB 14; Length 431;
Best Local Similarity 99.8%; Pred. No. 1.9e-200;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRALLARLLLCVIVVSDSKSGNELHQVPSNCDCLNGGTCVSNKXFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVVSDSKSGNELHQVPSNCDCLNGGTCVSNKXFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
QY 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
DB 121 YCNPDPNRRRPMCVYVGLKPLVOECMVHDCADGKSSPPEELKFOCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFLDYPKKEDYIYVLG 240
```

Db 181 GGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLMSPCWVISAHCFIDYPKKEDYIVYL 240
QY 241 RSRINSTOGEEMKEVENLILHKYSDTLAHNDIALKIRSEKRCQAPSRITOTICL 300
Db 241 RSRINSTOGEEMKEVENLILHKYSDTLAHNDIALKIRSEKRCQAPSRITOTICL 300
QY 301 PSWYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
Db 301 PSWYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCCGDSGGLVCSLQGRMTLTGIVSWRGALDKDXGVYTRVSHFLPWIR 420
Db 361 CAADPQWKTDSCCGDSGGLVCSLQGRMTLTGIVSWRGALDKDXGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
Db 421 SHTKEENGIAL 431

RESULT 10

US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.PI
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match 99.2%; Score 2376; DB 9; Length 431;

Best Local Similarity 99.5%; Pred. No. 3.9e-199;
Matches 429; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVLLVSDSKSNEHLQVPSNCDCLANGGTCVSNKYF--SNIMHCNCPKFKG 60
Db 1 MRALLARLLLCVLLVSDSKSNEHLQVPSNCDCLANGGTCVSNKYF--SNIMHCNCPKFKG 60
QY 61 HCEIDSKTYCEGNGHFYRGKASTDTWGRPCLPMNSATVTLQOTYHAHRSDALQLGLGKH 120
Db 61 HCEIDSKTYCEGNGHFYRGKASTDTWGRPCLPMNSATVTLQOTYHAHRSDALQLGLGKH 120
QY 121 YCRNPDRRRRWCVYVGLKPLVQECWVHDCADGKSSPPEELKFCGCGKTLARPKKII 180
Db 121 YCRNPDRRRRWCVYVGLKPLVQECWVHDCADGKSSPPEELKFCGCGKTLARPKKII 180

QY 181 GGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWVISAHCFIDYPKKEDYIVYL 240
Db 181 GGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWVISAHCFIDYPKKEDYIVYL 240
QY 241 RSRINSTOGEEMKEVENLILHKYSDTLAHNDIALKIRSEKRCQAPSRITOTICL 300
Db 241 RSRINSTOGEEMKEVENLILHKYSDTLAHNDIALKIRSEKRCQAPSRITOTICL 300
QY 301 PSWYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
Db 301 PSWYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCCGDSGGLVCSLQGRMTLTGIVSWRGALDKDXGVYTRVSHFLPWIR 420
Db 361 CAADPQWKTDSCCGDSGGLVCSLQGRMTLTGIVSWRGALDKDXGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
Db 421 SHTKEENGIAL 431

RESULT 11

US-10-360-101-266
; Sequence 266, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; TITLE OF INVENTION: Export and modification of (poly) peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 266
; LENGTH: 445
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of urokinase
US-10-360-101-266

Query Match 96.8%; Score 2317; DB 15; Length 445;

Best Local Similarity 96.9%; Pred. No. 5.9e-194;
Matches 431; Conservative 0; Mismatches 0; Indels 14; Gaps 7;

QY 1 MRALLARLLLCVLLVSDSKSNEHLQVPSNCDCLANGGTCVSNKYF--SNIMHCNCPKFKG 58
Db 1 MRALLARLLLCVLLVSDSKSNEHLQVPSNCDCLANGGTCVSNKYF--SNIMHCNCPKFKG 60
QY 59 GQHCIDSKTYCEGNGHFYRGKASTDTWGRPCLPMNSATVTLQOTY--HAHRSDALQLGL 116
Db 61 GQHCIDSKTYCEGNGHFYRGKASTDTWGRPCLPMNSATVTLQOTY--HAHRSDALQLGL 120
QY 117 GKYNCRNPNRRRPPCVYVGLKPLVQECWVHDCADGKSSPPEELKFCGCGKTLAR 174
Db 121 GKYNCRNPNRRRPPCVYVGLKPLVQECWVHDCADGKSSPPEELKFCGCGKTLAR 180
QY 175 PRFKIIGGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWVISA--THCFIDYPKK 232
Db 181 PRFKIIGGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWVISA--THCFIDYPKK 240
QY 233 EDYIVYIGSRRLNSNTGEMKEFEVENLILHKYSDTLAHNDIALKIRSEKRCQ 230
Db 241 EDYIVYIGSRRLNSNTGEMKEFEVENLILHKYSDTLAHNDIALKIRSEKRCQ 300
QY 291 PSRTIOTICLPSWYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKL--SHRECOOP 348
Db 301 PSRTIOTICLPSWYNDPOFGTSCETITFGKENSTDYLYPEQLKMTVVKL--SHRECOOP 360

QY 349 HYGSEVTTMKLCAADPQWKTDSCQDSGGELVCSLQGRMTLTGIVSWGRG--CALDKP 406
DB 361 HYGSEVTTMKLCAADPQWKTDSCQDSGGELVCSLQGRMTLTGIVSWGRGFTCALDKP 420
QY 407 GYTVRSHFLPMIRSHKTEENGAL 431
DB 421 GYTVRSHFLPMIRSHKTEENGAL 445

RESULT 12

US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRM
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 96.1%; Score 2301; DB 9; Length 411;
Best Local Similarity 100.0%; Pred. No. 1,3e-192;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SNEHQVPSNCDCLNGSTCVSNKYPSNIHWCNCPKFGQHCETDSEKTCYEGNGHFRYG 80
DB 1 SNEHQVPSNCDCLNGSTCVSNKYPSNIHWCNCPKFGQHCETDSEKTCYEGNGHFRYG 60
QY 81 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYGLK 140
DB 61 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYGLK 120
QY 141 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENOQWFAIYRRH 200
DB 121 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENOQWFAIYRRH 180
QY 201 RGGSVTVVCGSLISPCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQEMKFEVENLI 260
DB 181 RGGSVTVVCGSLISPCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQEMKFEVENLI 240
QY 261 LHKDYADTLAHHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFSGCEITGFGK 320
DB 241 LHKDYADTLAHHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFSGCEITGFGK 300
QY 321 ENSTDVLYPEQLKMTYVKLISHRECOOPHYGSEVTTMKLCAADPQWKTDSCQDSGGPL 380
DB 301 ENSTDVLYPEQLKMTYVKLISHRECOOPHYGSEVTTMKLCAADPQWKTDSCQDSGGPL 360
QY 381 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 431
DB 361 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 411

RESULT 13

US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRAACYCLINE-INDUCED PLEURITIS IN RABBITS

; FILE REFERENCE: UTSN:02205
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRM
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 95.7%; Score 2291; DB 15; Length 411;
Best Local Similarity 99.8%; Pred. No. 1e-191;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SNEHQVPSNCDCLNGSTCVSNKYPSNIHWCNCPKFGQHCETDSEKTCYEGNGHFRYG 80
DB 1 SNEHQVPSNCDCLNGSTCVSNKYPSNIHWCNCPKFGQHCETDSEKTCYEGNGHFRYG 60
QY 81 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYGLK 140
DB 61 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYGLK 120
QY 141 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENOQWFAIYRRH 200
DB 121 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENOQWFAIYRRH 180
QY 201 RGGSVTVVCGSLISPCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQEMKFEVENLI 260
DB 181 RGGSVTVVCGSLISPCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQEMKFEVENLI 240
QY 261 LHKDYADTLAHHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFSGCEITGFGK 320
DB 241 LHKDYADTLAHHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFSGCEITGFGK 300
QY 321 ENSTDVLYPEQLKMTYVKLISHRECOOPHYGSEVTTMKLCAADPQWKTDSCQDSGGPL 380
DB 301 ENSTDVLYPEQLKMTYVKLISHRECOOPHYGSEVTTMKLCAADPQWKTDSCQDSGGPL 360
QY 381 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 431
DB 361 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 411

RESULT 14

US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRM
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match 93.7%; Score 2243; DB 9; Length 403;
Best Local Similarity 98.1%; Pred. No. 1.5e-187;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

QY 21 SNEIHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKFGGQHCEDKSKTCYEGNGHFRG 80
Db 1 SNEIHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKFGGQHCEDKSKTCYEGNGHFRG 60
QY 81 KASDTMGRPCLPNNSATVLOQTYHAHRSDALQGLGKHNCRPNDRNRPRWCYVQVGLK 140
Db 61 KASDTMGRPCLPNNSATVLOQTYHAHRSDALQGLGKHNCRPNDRNRPRWCYVQVGLK 120
QY 141 PLYOECWHDCAADGKPPPEELKFGGQKTLRPRFKTIGSEFTTIENQWPAIYRRH 200
Db 121 PLYOECWHDCAADGK-----LKFQCGQKTLRPRFKTIGSEFTTIENQWPAIYRRH 172
QY 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENLI 260
Db 173 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENLI 232
QY 261 LHKDYSADTLAHNDIALKLRKSGRCAPSRITQITCLPSMYNDPQFTSCETITGFGK 320
Db 233 LHKDYSADTLAHNDIALKLRKSGRCAPSRITQITCLPSMYNDPQFTSCETITGFGK 292
QY 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPOMKTDSCGDSGGPL 380
Db 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPOMKTDSCGDSGGPL 352
QY 381 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
Db 353 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

```

RESULT 15

```

US-09-880-503-7
; Sequence 7, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212, 847
; PRIORITY DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-7

```

Query Match 72.6%; Score 1737; DB 9; Length 323;

Best Local Similarity 78.6%; Pred. No. 1.9e-143; Indels 88; Gaps 1;

Matches 323; Conservative 0; Mismatches 0; Indels 88; Gaps 1;

```

QY 21 SNEIHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKFGGQHCEDKSKTCYEGNGHFRG 80
Db 1 SNEIHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKFGGQHCEDKSKTCYEGNGHFRG 47
QY 81 KASDTMGRPCLPNNSATVLOQTYHAHRSDALQGLGKHNCRPNDRNRPRWCYVQVGLK 140
Db 48 ----- 47
QY 141 PLYOECWHDCAADGKPPPEELKFGGQKTLRPRFKTIGSEFTTIENQWPAIYRRH 200
Db 48 -----KSPPEELKFGGQKTLRPRFKTIGSEFTTIENQWPAIYRRH 92
QY 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENLI 260
Db 93 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENLI 152
QY 261 LHKDYSADTLAHNDIALKLRKSGRCAPSRITQITCLPSMYNDPQFTSCETITGFGK 320

```

```

Db 153 LHKDYSADTLAHNDIALKLRKSGRCAPSRITQITCLPSMYNDPQFTSCETITGFGK 212
QY 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPOMKTDSCGDSGGPL 380
Db 213 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPOMKTDSCGDSGGPL 272
QY 381 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
Db 273 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 323

```

Search completed: March 18, 2004, 11:36:07
Job time: 129.513 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 44.326 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-2

Perfect score: 2394
Sequence: 1 MRLALRLLCVLVSDSKG.....VSHFLPWRSHYKENGAL 431

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 78:*
2: p1r1:*
3: p1r2:*
4: p1r3:*
5: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2391	99.9	431	1 UKHU	u-plasminogen acti
2	2228	93.1	433	1 UKBAY	u-plasminogen acti
3	1928.5	80.6	442	1 UKPG	u-plasminogen acti
4	1835	76.6	433	1 JN0560	u-plasminogen acti
5	1728.5	72.2	432	1 S18932	u-plasminogen acti
6	1700	71.0	433	1 UKMS	u-plasminogen acti
7	1009	42.1	434	1 A35005	u-plasminogen acti
8	868.5	36.3	477	2 JS0598	u-plasminogen acti
9	868.5	36.3	562	1 UKHUT	u-plasminogen acti
10	867.5	36.2	431	2 JS0599	u-plasminogen acti
11	867.5	36.2	477	1 A34369	u-plasminogen acti
12	847.5	35.4	559	1 A35029	u-plasminogen acti
13	841.5	35.2	477	2 JS0597	u-plasminogen acti
14	833.5	34.8	559	1 A29941	u-plasminogen acti
15	752	31.4	394	2 JS0600	u-plasminogen acti
16	735.5	30.7	655	1 A46688	hepatocyte growth
17	718.5	30.0	603	2 S28941	coagulation factor
18	692	28.9	615	1 JF0178	coagulation factor
19	674.5	28.2	558	2 JCS578	plasma hyaluronan
20	661.5	27.6	560	1 UC4795	plasma hyaluronan
21	643.5	26.9	553	2 S45281	coagulation factor
22	507.5	21.2	460	2 B61545	plasmin (EC 3.4.21
23	504.5	21.1	790	1 PLUG	plasmin (EC 3.4.21
24	501	20.9	810	1 PLUH	plasmin (EC 3.4.21
25	497	20.8	810	2 B30848	plasmin (EC 3.4.21
26	484.5	20.2	812	1 PLMS	plasmin (EC 3.4.21
27	484	20.2	1420	2 A32869	apoptoprotein(a)
28	474	19.8	4548	1 S00657	apoptoprotein(a) (EC

30	457.5	19.1	455	2 A61545	plasmin (EC 3.4.21
31	445	18.6	761	2 JC5759	brain-specific ser
32	435	18.2	810	2 I46260	plasmin (EC 3.4.21
33	410	17.1	638	1 K0MSPL	plasmin (EC 3.4.21
34	408.5	17.1	343	1 A57014	proctasin (EC 3.4.
35	406	17.0	711	1 A47136	macrophage-stimula
36	402.5	16.8	417	1 S00845	hepsin (EC 3.4.21.
37	402	16.8	638	1 K0HUP	hepsin (EC 3.4.21.
38	401	16.8	416	1 S33777	membrane-bound arg
39	401	16.8	855	2 JC7731	trypsin (EC 3.4.21
40	395.5	16.5	248	2 S5066	trypsin (EC 3.4.21
41	394.5	16.5	229	1 TRB0TR	trypsin (EC 3.4.21
42	394.5	16.5	716	1 JC5061	macrophage-stimula
43	393	16.4	728	1 A40332	hepatocyte growth
44	392	16.4	716	1 A40332	macrophage-stimula
45	390.5	16.3	247	2 S13813	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1

UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
in form

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1982 #sequence_revision 04-Dec-1986 #text change 15-Sep-2000

C;Accession: A00931; I52209; J0102; A37561; I38102; S65783; A37562; A37563; A37564; A356

R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boate, S.; Blaai, F.

Nucleic Acids Res. 13, 2759-2771, 1985

A;Title: The human urokinase-plasminogen activator gene and its promoter.

A;Reference number: A00931; M01D:85215647; PMID:2987867

A;Accession: A00931

A;Molecule type: DNA

A;Residues: 1-431 <RC>

A;Cross-references: GB:K03429; NID:937601; PIDN:CAA26268.1; PID:G1834524

A;Note: the authors translated the codon ATG for residue 214 as Ile

R;Nagamine, Y.; Pearson, D.; Gratian, M.

Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine

A;Reference number: I52209; M01D:86050639; PMID:3933505

A;Accession: I52209

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 145-161 <NAG1>

A;Cross-references: GB:K03427; NID:9340174; PIDN:AAA61257.1; PID:9340175

R;Nagai, M.; Hiratake, R.; Kaneda, T.; Hayasaka, N.; Nishida, M.; Suyama, J

Gene 36, 183-188, 1985

A;Title: Molecular cloning of cDNA coding for human preprourokinase.

A;Reference number: J0102; M01D:86056954; PMID:2415429

A;Accession: J0102

A;Molecule type: mRNA

A;Residues: 1-213, 'I', 215-431 <NAG2>

A;Cross-references: GB:K03426; NID:9340155; PIDN:AA971138.1; PID:9340158; GB:D00244; NID

R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F.

Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984

A;Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA

A;Reference number: A37561; M01D:84272706; PMID:6589620

A;Accession: A37561

A;Molecule type: mRNA

A;Residues: 66-431 <VER>

A;Cross-references: GB:D00244; NID:9220138

R;Jacobs, P.; Crivador, A.; Lottia, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen,

DNA 4, 139-146, 1985

A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prepr

A;Reference number: I38102; M01D:85203359; PMID:3888571

A;Accession: I38102

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>

A;Cross-references: EMBL:X02760; NID:935297; PIDN:CAA26535.1; PID:935298

R; Yoshimoto, M.; Uehiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
A:Title: Characterization of single chain urokinase-type plasminogen activator with a no
A:Reference number: 965783; MUID:96186279; PMID:8652631
A:Accession: S65783
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 21-140, 'L', 142-213, 'T', 215-431 <YOS>
A:Cross-references: EMBL:D11143; NID:91311467; PIDN:BA01919.1; PID:g119928
R; Gunzler, W.A.; Steffens, G.J.; Oetting, F.; Kilm, S.M.A.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A:Title: The primary structure of high molecular mass urokinase from human urine.
A:Reference number: A37562; MUID:83055084; PMID:6754569
A:Accession: A37562
A:Molecule type: protein
A:Residues: 21-177 <GUN>
R; Schaller, U.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergler, W.; Studer, R.O.
Eur. J. Biochem. 125, 251-257, 1982
A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and preli
A:Reference number: A37563; MUID:83003608; PMID:6749491
A:Accession: A37563
A:Molecule type: protein
A:Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>
R; Steffens, G.J.; Gunzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A:Title: The complete amino acid sequence of low molecular mass urokinase from human ur
A:Reference number: A37564; MUID:83055099; PMID:6754572
A:Accession: A37564
A:Molecule type: protein
A:Residues: 158-410 <STE>
R; Kautzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
Biochem. Biophys. Res. Commun. 171, 401-406, 1990
A:Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinat
A:Reference number: A35689; MUID:90365737; PMID:2393398
A:Accession: A35689
A:Molecule type: protein
A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
A:Note: identification of a fucose and attempt to determine its attachment site
R; Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazur, A.; Henkin, J.; Goltz
Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
A:Reference number: A36697; MUID:91097529; PMID:2125213
A:Accession: A36697
A:Molecule type: protein
A:Residues: 21-34 <RAB>
R; Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
submitted to the Brookhaven Protein Data Bank, July 1993
A:Reference number: A51255; PDB:1KDU
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R; Li, X.; Smith, R.A.G.; Dobson, C.M.
Biochemistry 31, 9562-9571, 1992
A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
A:Reference number: A44375; MUID:93003110; PMID:1327118
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
submitted to the Brookhaven Protein Data Bank, January 1994
A:Reference number: A66822; PDB:1URK
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
submitted to the Brookhaven Protein Data Bank, July 1995
A:Reference number: A66058; PDB:1LWM
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
A:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
C:Genetics: GDB:PLAU
A:Gene: GDB:119497; OMIM:191840
A:Cross-references: GDB:10924-10924
A:Map position: 10924-10924
A:Introns: 19/3, 25/1, 65/1, 123/2, 154/1, 227/2, 277/1, 324/1, 373/3
C:Function: proteolytically activates plasminogen
A:Description: proteolytically activates plasminogen
A:Pathway: fibrinolysis
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteinat
F:1-20/Domain: signal sequence #status predicted <SIG>
F:1-43/Product: urokinase-type plasminogen activator, single chain form #status predict
F:21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MPA>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KR>
F:156-177/Product: urokinase-type plasminogen activator chain B #status experimental <MP
F:179-419/Domain: trypsin homology <TRY>
F:31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-400
F:38/Binding site: carboxylate (Thr) (covalent) #status predicted
F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F:224, 275, 376/Active site: His, Asp, Ser #status experimental
F:322/Binding site: carboxylate (Asn) (covalent) #status experimental
Query Match 99.8%; Score 2391; DB 1; Length 431;
Best Local Similarity 99.8%; Pred. No. 2, 3e-175;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1 MRALLARLLLCVLLVSDSKGSNELHQPNSCDLNGTCVSNKYPSNHNWNCPPKFGQ 60
1 MRALLARLLLCVLLVSDSKGSNELHQPNSCDLNGTCVSNKYPSNHNWNCPPKFGQ 60
Qy HCEIDSKTCYEGNGHFGYKASTDTPMGRPCPMNSATYLQOTYHNRSDALQGLGKN 120
HCEIDSKTCYEGNGHFGYKASTDTPMGRPCPMNSATYLQOTYHNRSDALQGLGKN 120
Db HCEIDSKTCYEGNGHFGYKASTDTPMGRPCPMNSATYLQOTYHNRSDALQGLGKN 120
Qy YCNPNRRPNCYVGVGLKPLVQECMVNDGDKGKSSPPEELKFCQCKTLRPFKII 180
YCNPNRRPNCYVGVGLKPLVQECMVNDGDKGKSSPPEELKFCQCKTLRPFKII 180
Db YCNPNRRPNCYVGVGLKPLVQECMVNDGDKGKSSPPEELKFCQCKTLRPFKII 180
Qy GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLSCPCWYISNHCITDYPKKEDYVYIG 240
GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLSCPCWYISNHCITDYPKKEDYVYIG 240
Db GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLSCPCWYISNHCITDYPKKEDYVYIG 240
Qy RSLNSTQCEMPEVENLILHDOVSADTLAHNDLALKIRKEKRCAPSTTIQICL 300
RSLNSTQCEMPEVENLILHDOVSADTLAHNDLALKIRKEKRCAPSTTIQICL 300
Db RSLNSTQCEMPEVENLILHDOVSADTLAHNDLALKIRKEKRCAPSTTIQICL 300
Qy PSMYNDPQFQSCIEITFGKENSTDVLYPEQLKMTYVKLISHRECOQPHYGSEVTTKML 360
PSMYNDPQFQSCIEITFGKENSTDVLYPEQLKMTYVKLISHRECOQPHYGSEVTTKML 360
Db PSMYNDPQFQSCIEITFGKENSTDVLYPEQLKMTYVKLISHRECOQPHYGSEVTTKML 360
Qy CAADPQWKTDSCGDSGGLVCSLOGRMITLTVISWGRGALKDKPQVTVRVSHFLPWIR 420
CAADPQWKTDSCGDSGGLVCSLOGRMITLTVISWGRGALKDKPQVTVRVSHFLPWIR 420
Db CAADPQWKTDSCGDSGGLVCSLOGRMITLTVISWGRGALKDKPQVTVRVSHFLPWIR 420
Qy SHTEKENGIAL 431
SHTEKENGIAL 431
Db SHTEKENGIAL 431
RESULT 2
UKBAY
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C:Date: 31-Dec-1991 #sequence, revision 31-Dec-1991 #text, change 18-Jun-1999
C:Accession: S14687; MUID:90287734; PMID:2113276
A:Reference number: S14687
A:Accession: S14687
A:Molecule type: mRNA
A:Residues: 1-433 <AYU>
A:Cross-references: EMBL:X51935; NID:938130; PIDN:CA36200.1; PID:g38131
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-176/Product: plasminogen activator chain A #status predicted <ACH>
F:30-61/Domain: EGF homology <EGF>
F:169-150/Domain: kringle homology <KR>
F:178-433/Product: plasminogen activator chain B #status predicted <BCH>

F:178-421/Domain: trypsin homology <TRY>
 F:167-298,208-224,216-287,315-384,374-402/Disulfide bonds: #status predicted
 F:223,274,378/Active site: His, Asp, Ser #status predicted
 F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.1%; Score 2228; DB 1; Length 433;
 Best Local Similarity 92.4%; Pred. No. 6,9e-153;
 Matches 401; Conservative 17; Mismatches 12; Indels 4; Gaps 2;

```

QY 1 MRLLARLLLCVAVSDSKSGNELHGV--DSNCDCLNGGTCVSNKYPFNIHMCNCPKPKGG 60
DB 1 MRLLARLLLCVAVSDSKSGNELHGV--DSNCDCLNGGTCVSNKYPFNIHMCNCPKPKGG 59
QY 61 HCEIDSKSKTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLAQTYHAHSDALQGLGKN 120
DB 60 HCEIDSKSKTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLAQTYHAHSDALQGLGKN 119
QY 121 YCRNPDRRRPWCYVQVGLKPLVQECVHNCADGKSSPPEELKFCGQKTLRPRKXII 180
DB 120 YCRNPDRRRPWCYVQVGLKPLVQECVHNCADGKSSPPEELKFCGQKTLRPRKXII 179
QY 181 GGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240
DB 180 GGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 239
QY 241 RSLNSNTQGEKFEVENILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICL 300
DB 240 RSLNSNTQGEKFEVENILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICL 299
QY 301 PSNTYNDPQ--FGTSCITTFGKENSVDLYLPBOLKMTVVKLISHRECOOPHYGSEVTT 357
DB 300 PSNTYNDPNDPFGTSCITTFGKENSVDLYLPBOLKMTVVKLISHRECOOPHYGSEVTT 359
QY 358 KMLCAADPQWKTDSCQSDSGPLVCSLQGRMTLTGIYSWGRGALKDKPQVTVRVSHFLP 417
DB 360 KMLCAADPQWKTDSCQSDSGPLVCSLQGRMTLTGIYSWGRGALKDKPQVTVRVSHFLP 419
QY 418 WIRSHTEENGIAL 431
DB 420 WIRSHTEENGIAL 433

```

RESULT 3

UNPKG
 u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N:Alternate names: uPA
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R:Nagamine, Y.; Pearson, D.; Altun, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; M01D:85087954; PMID:6096832
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <MAG1>
 A:Experimental source: kidney cell line LLC-PK1
 R:Nagamine, Y.
 submitted to the Protein Sequence Database, December 1986
 A:Reference number: A37566
 A:Contents: annotation; correction to residue 241
 C:Gene(s):
 A:Insertions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:23-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:190-430/Domain: trypsin homology <TRY>
 F:190-430/Domain: trypsin homology <TRY>
 F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted <BGH>
 F:179-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted

F:235,286,387/Active site: His, Asp, Ser #status predicted
 Query Match 80.6%; Score 1928.5; DB 1; Length 442;
 Best Local Similarity 79.4%; Pred. No. 5,8e-140;
 Matches 350; Conservative 33; Mismatches 47; Indels 11; Gaps 2;

```

QY 1 MRLLARLLLCVAVSDSKSGNELHGV--DSNCDCLNGGTCVSNKYPFNIHMCNCPKPKGG 58
DB 1 MRLLARLLLCVAVSDSKSGNELHGV--DSNCDCLNGGTCVSNKYPFNIHMCNCPKPKGG 60
QY 59 GCHCEIDSKSKTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLAQTYHAHSDALQGLGK 118
DB 61 GCHCEIDSKSKTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLAQTYHAHSDALQGLGK 120
QY 119 HNYCRNPDRRRPWCYVQVGLKPLVQECVHNCADGKSSPPEELKFCGQKTLRPRKXII 169
DB 121 HNYCRNPDRRRPWCYVQVGLKPLVQECVHNCADGKSSPPEELKFCGQKTLRPRKXII 180
QY 170 QKTLRPRKXIIIGGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISAHCFIDY 229
DB 181 QKTLRPRKXIIIGGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISAHCFIDY 240
QY 230 PKKEDYIYVGRSLNSNTQGEKFEVENILHKDYSADTLAHNDIALKIRSKGRCA 289
DB 241 QKEDYIYVGRSLNSNTQGEKFEVENILHKDYSADTLAHNDIALKIRSKGRCA 300
QY 290 QPSRTIQTICLPQMYNDPQGTSCITTFGKENSVDLYLPBOLKMTVVKLISHRECOOPH 349
DB 301 QPSRTIQTICLPQMYNDPQGTSCITTFGKENSVDLYLPBOLKMTVVKLISHRECOOPH 360
QY 350 YGSEVTTKMLCAADPQWKTDSCQSDSGPLVCSLQGRMTLTGIYSWGRGALKDKPQVTVRVSHFLP 409
DB 361 YGSEVTTKMLCAADPQWKTDSCQSDSGPLVCSLQGRMTLTGIYSWGRGALKDKPQVTVRVSHFLP 420
QY 410 TRYSHFLPTRYSHTEENGIAL 430
DB 421 TRYSHFLPTRYSHTEENGIAL 441

```

RESULT 4

UNPKG
 u-plasminogen activator (EC 3.4.21.73) precursor - bovine
 N:Alternate names: uPA
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: JN0560
 R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schlenning, W.D.
 Gene 125, 177-183, 1993
 A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
 A:Reference number: JN0560; M01D:93216119; PMID:8385052
 A:Accession: JN0560
 A:Molecule type: mRNA
 A:Residues: 1-433 <KRA>
 A:Cross-references: GB:103546; NID:G163800; PIDN:AA51419.1; PID:G163801
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
 F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F:181-421/Domain: trypsin homology <TRY>
 F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.6%; Score 1835; DB 1; Length 433;
 Best Local Similarity 75.1%; Pred. No. 8e-133;
 Matches 325; Conservative 46; Mismatches 60; Indels 2; Gaps 1;

```

QY 1 MRLLARLLLCVAVSDSKSGNELHGV--DSNCDCLNGGTCVSNKYPFNIHMCNCPKPKGG 58
DB 1 MRLLARLLLCVAVSDSKSGNELHGV--DSNCDCLNGGTCVSNKYPFNIHMCNCPKPKGG 60

```


[illegible]

RESULT 5
518932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N/Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C/Species: Rattus norvegicus (Norway rat)
C/Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
C/Accession: S24604; I60186; I53472; S18932
R/Rabhanl, S.A.
submitted to the EMBL Data Library, April 1992
A/Reference number: S24604
A/Accession: S24604
A/Molecule type: mRNA
A/Residues: 1-15, 'H', '17-23', 'G', '25-331', 'N', '333-432 <RAB>
A/Cross-references: EMBL:X65651; NID:G57456; PIDN:CAA4601.1; PID:G57457
A/Experimental source: tissue kidney
R/Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A/Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A/Reference number: I60186; MUID:92233409; PMID:1568219
A/Accession: I60186
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-432 <RES>
A/Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466
A/Experimental source: Strain Fischer 344; tissue mammary
R/Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasl, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A/Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A/Reference number: I53472; MUID:9233549; PMID:1321734
A/Accession: I53472
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 31-62 <RE2>
A/Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279
C/Genetics:
A/Gene: uPA
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C/Keywords: glycoprotein, heterodimer, hydrolase, kringle, serine proteinase
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F/31-62/Domain: EGF homology <EGF>
F/70-151/Domain: kringle homology <KR>
F/179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

```

F:179-320/Domains: trypsin homology <TRY>
F:166-300,210-226,218-289,314-303,146-162,373-401/Dsulfide bonds: #status predicted
F:1225,276,377/Active site: His, Asp, Ser #status predicted

Query Match      72.2%      Score 1228.5;   DB: 1;   Length 432;
Best Local Similarity 71.3%      Pred. No. 11e-124;
Matches 308; Conservative 51; Mismatches 70; Indels 3; Gaps 3;

```

QY	1	MRALLATLLLCVLYVSSDKGSNLEH--QVPSNCDCLNGCTCVSKKPSYNSIHMCNCPKKFG	59
Db	1	MRVWLASLFICAL-VANSBGSELEIASDESNCGCGNGVCVSTKYSIRSCCPKKFG	59
QY	60	QHCIEDSKCTCYEGSNGHFRGKASTDIMGRCPLPWNSATVLOOTYAHRESDALQGLGKH	119
Db	60	EHCEIDTSKCTCYHNGSYRGKANTDTKGRPLCLAMNSPAVLQGTYNAHRESDALSLGKH	119
QY	120	NYCNPENRRRPMCTYVOVGLKPLVQEGMVDCAADGKKPSSPBEELKQCCQKTLRRPKFI	179
Db	120	NYCNPENRRRPMCTYVOVGLKPLVQEGMVDCAADGKKPSSPBEELKQCCQKTLRRPKFI	179
QY	180	IGGEFTTIENQWPEAIIYRRHRGGS--VTYYCGGSLISPCWVISATHCFIDYPPKEDYIV	238
Db	180	VGGEFTYVENQWPEAIIYLRKKGSPSPFCGGSLSISPCWVAATHCFVNGPKKEEVIV	239
QY	239	LGRSLNSNTQGEKMEFVENLIIHKDYSADTLAHNDIALILKTRSGEGRCAQPSRTIOTT	298
Db	240	LGSGRKNSYNPGEKKEFEVEQLIIHEDFSDETLFHHNDIALILKTRSTGQCAQPSRTIOTT	299
QY	299	CLPEMNVDPQEGTSCETLGRGKENSVDYLAPBOLXMTVYVLTISHRCQOQHYYGSEVTRK	358
Db	300	CLPEPFEDAPPGSCELTGFGQESADIFYFKOLKMSVYVLTISHBQCKQPHYYGSELTNYK	359
QY	359	MLCAADPQWKTIDSQCGDSGGPLVCSLQGRVTLTLGIYSWGRGALKDKPGVYTVSFLLPW	418
Db	360	MLCAADPEMKTIDSCGDSGGPLICNIDGRFTLSGIYSWGSGCAEKQKPGVYTVSFLLPW	419
QY	419	IRSHTKENGLA 430	
Db	420	IQSHIGENGLA 431	

RESULT 6

UKMS

u-plasminogen activator (EC 3.4.21.73) precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #ext_change 18-Jun-1999

C:Accession: A29420; A24615

R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.

Biochemistry 26, 8270-8279, 1987

A>Title: The murine urokinase-type plasminogen activator gene.

A:Reference number: A29420; MUID:86163489; PMID:2831940

A:Accession: A29420

A:Molecule type: DNA

A:Releases: 1-433 <DEG>

A:Cross-references: GB:I19722; NID:g202296; PIDN:AAA0539.1; PID:g202297

R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godéau, F.; Nagamine, Y.; Reich, E.; Kocher, E.

Jur J. Biochem. 148, 225-232, 1985

A>Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase

A:Reference number: A24615; MUID:85179474; PMID:2285383

A:Accession: A24615

A:Molecule type: mRNA

A:Releases: 1-433 <BEL>

A:Cross-references: GB:X02389; NID:g55127; PIDN:CMA6231.1; PID:g55128

C:Genetics:

A:Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringlike homology; trykingle homology; trykingle homology; trykingle homology <TRY>

C:Keywords: glycoprotein; heterodimer; hydrolase; kringlike; serine proteinase

F:1-20/Domain: signal sequence #status predicted <Sig>

F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F:32-63/Domain: EGF homology <EGF>

F:71-152/Domain: kringlike homology <KRK>

F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F:180-421/Domain: trypsin homology <TRY>

F:169-301,211-227,219-290,315-384,347-363,374-402/disulfide bonds: #status predicted
F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 71.0%; Score 1700; DB 1; Length 433;
Best Local Similarity 69.2%; Pred. No. 1,7e-122;
Matches 299; Conservative 56; Mismatches 75; Indels 2; Gaps 2;

QY 1 MRALLARLLCVLVSDSKSGSNELHGV-PSNCDLNGTCVSNKYFSNIHMCNPKKRG 59
DB 1 MKWLAFLFCALVVKNSBEGSVLGAPDESNCGONGGVCSYKFSIRRCSCPRKFG 60
QY 60 OHCEIDSKTCYBENGHFYRGKASTDPMGRPLPMNSATVLOQTYHAHRSALDGLGKH 119
DB 61 EHCIDASKTCYHENGSGSYRGKANTDTGRCLAMNAPAVLQKPYNAHPPAISLGIGKH 120
QY 120 NYSRNPNNRRPWCYVGVGLKPLVQECVHDCADKKPSPPEELKFOCGQTLRPFKI 179
DB 121 NYSRNPNNRRPWCYVGVGLKPLVQECVHDCADKKPSPPEELKFOCGQTLRPFKI 180
QY 180 TGEFTTIENQPMFAIYRRHGGV-VTVYCGSLISPCWVISAHCFIDYPKKEDYIV 238
DB 181 VGEFTTEVENQPMFAIYQKNGGSPSPKCGSLISPCWVISAHCFIDYPKKEDYIV 240
QY 239 LGSRLNSNTQGEKFEVENLILHKDYASDTLHNDIALKIRSKGRCAPSRITQI 298
DB 241 LGSRLNSNTQGEKFEVENLILHKDYASDTLHNDIALKIRSKGRCAPSRITQI 300
QY 299 CLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 358
DB 301 CLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 360
QY 359 MLCADPQMTKDSQCGSGGGLVCSLQGRMTLTGIVSGRCALDKRQVYTRVSHPLPW 418
DB 361 MLCADPQMTKDSQCGSGGGLVCSLQGRMTLTGIVSGRCALDKRQVYTRVSHPLPW 420
QY 419 IRSHTEENGLA 430
DB 421 IOSHTEENGLA 432

RESULT 7

A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N/Alternate names: uPA
C/Species: Gallus gallus (chicken)
C/Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C/Accession: A35005
R/Residue: N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A/Title: The chicken urokinase-type plasminogen activator gene.
A/Reference number: A35005; MUID:90110185; PMID:2295632
A/Accession: A35005
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-434 <ES>
A/Cross-references: GB:005187; NID:g212858; PIDN:AAA49131.1; PID:g212859
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-111/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:40-71/Domain: EGF homology <EGF>
F:79-158/Domain: kringle homology <KR>
F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:173-416/Domain: trypsin homology <TRY>
F:162-296,202-218,210-285,310-379,342-350,369-397/Disulfide bonds: #status predicted
F:217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 42.1%; Score 1009; DB 1; Length 434;
Best Local Similarity 46.0%; Pred. No. 1,2e-69;
Matches 196; Conservative 66; Mismatches 134; Indels 30; Gaps 8;
QY 10 LCVLVVS-DS-----KGSNEHGVPSNCDLNGTCVSNKYFSNIHMCNPKKRG 61

DB 11 LCVLVGSDVYIRQYKLSHKHRRPDRBEQCLNGTCITTYRFSQIKRCLCPREGGLH 70
QY 62 CEIDSKTCYBENGHFYRGKASTDPMGRPLPMNSATVLOQTYHAHRSALDGLGKH 120
DB 71 CEIDTNSICTSGNEDIRGMAEDP-----GCLYWHPSVIRRGVDYHADIKNALDGLGKH 126
QY 121 YCRNPNNRRPWCYVGVGLKPLVQECVHDCADKKPSPPEELKFOCGQTLRPFKI 180
DB 127 YCRNPNNRRPWCYTK--RRYSIOE-----TCSSTIEKERTCGQSFSGYFKIV 174
QY 181 GGEFTTIENQPMFAIYRRHGGV-VTVYCGSLISPCWVISAHCFIDYPKKEDYI 236
DB 175 GSGQAEVETPMFATGIFQNTM-GTDOPLCGSLIDPCWVILPAACFYNPTRKQPNKSVYK 233
QY 237 VYLSRLNSNTQGEKFEVENLILHKDYASDTLHNDIALKIRSKGRCAPSRITQ 296
DB 234 VFLGKSLTNNDEHQVFMWDEIISHDFDTHGNGNDLALIRIRASQCAVESYVR 293
QY 297 TILPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 356
DB 294 TILPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 353
QY 357 TKNLCADPQMTKDSQCGSGGGLVCSLQGRMTLTGIVSGRCALDKRQVYTRVSHPL 416
DB 354 DNWVCAQDPLMETDCKGDSGGVPCENGRMTLYGIVSWGDCAKKNGVYTRVRYL 413
QY 417 PMIRSH 422
DB 414 NWIDSN 419

RESULT 8

JS0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N/Alternate names: tissue plasminogen activator
C/Species: Desmodus rotundus (common vampire bat)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C/Accession: JS0598
R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dom
Gene 105, 229-237, 1991
A/Title: The plasminogen activator family from the salivary gland of the vampire bat Desm
A/Reference number: JS0597; MUID:92039036; PMID:1937019
A/Accession: JS0598
A/Molecule type: mRNA
A/Residues: 1-477 <KRA>
A/Cross-references: GB:M63988; NID:g166074; PIDN:AAA1593.1; PID:g166075
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F:47-79/Domain: fibronectin type I repeat homology <IFR>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KR>
F:226-471/Domain: trypsin homology <TRY>
F:47-79,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:185,398/Binding site: carbohydrate (Aan) (covalent) #status predicted
F:225-326/Cleavage site: His-Ser (plasmin) #status predicted
F:212,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.3%; Score 868.5; DB 2; Length 477;
Best Local Similarity 42.7%; Pred. No. 7,6e-59;
Matches 179; Conservative 61; Mismatches 150; Indels 29; Gaps 9;

QY 19 KGSNEHGVPSNCDLNGTCVSNKYFSNIHMCNPKKRG 74
DB 74 RGLAQCHTVPVKSCSELRGNGTCVQWASFSDF-VCOCPKGYTGKCEVDTHATCYKDQ 132
QY 75 GHYRGKASTDPMGRPLPMNSATVLOQTYHAHRSALDGLGKHNYCRPNRRPWCY 134
DB 133 GVTYRGKASTDPMGRPLPMNSATVLOQTYHAHRSALDGLGKHNYCRPNRRPWCY 132
QY 135 VQVGLKPLVQECVHDCADKKPSPPEELKFOCG-QKTLRPFKIIGGEFTTIENQPM 193

Query Match	36.3%	Score	868.5;	DB 1;	Length	562;
Best Local Similarity	37.9%	Pred	No.9.1e-59;			
Matches 187;	Conservative	57;	Mismatches	160;	Indels	89;
					Gaps	11

```

0Y 18 SKGSNELHQV-A-SUND---CLNGGTVCYNKXFSNIMHMCNCKKCGQHCCEIDSKCTYEG 73
Db 72 NSGRAQCHSVYKSCSEBRCFNGGTCQOALFSDP-VQCEPFGAGKCCCEIDTRATCYED 130
0Y 74 NGHFYRGA-SYDTMGRPCLPMNSATVLTQOTHAHRSDALQJLGKHYCRYPDNRBRPWC 133
Db 131 QGISYRGWSTABESGAECTMNSSALAQKPYSGRRPDALIRLGLGNHNYCRNPDSDKPMC 190
0Y 134 YVQVGLKELVQECWVHDCADG-----154
Db 191 YVFRAGKXSSFCGCPACSEGNDCOYFNGSAYRGTHSLTESGASCLPMNSMILIGKYYT 250
0Y 155 -KKPS-----SPEELK-----FOCG-QKTLRPRFK 178
Db 251 AQNSQAQALGLGKHNYCRNPDGDAKPMWCHVLKNRRLTWECYDVPSCSTGRLQXSPQFR 310
0Y 179 IIGSEFTTIEQPMFPAAYRRH-RGGSVTVYCGGSLSPCWIVISATHCIFDYPKKEDEYV 237
Db 311 IKGGIFADIASHPMQALFACHRRSPGERFLCGGILLISSCWLIAAHFOERPPHNLTV 370
0Y 238 YLGSRSLNSNTOGEMKFEVENTILKHODYSADTLAHNHDIALTKKRSKRGCAOCSRTIOT 297
Db 371 ILGRYRVYVPOSEBQKFEVEKTYIYHKEFDPT--YNDNIALDLQKSDSSRCAQESSVVRT 428
0Y 298 ICLPSMYNDPOFGTSCEITGFGKENSVDYLYPEQLKMTVLKLSHRECQPHYVGSVYTT 357
Db 429 VCLPRPADQLQDWMTECELSGYGKHIALSPFSERIKAEHVRLYPSRSRTSGHLNRRYTD 488
0Y 358 KMLCAAD-----PQWKT-DSCQDSDGSLVCSLQGRMTLTGIVSMGRGALKDXPGYYTR 411
Db 489 NMLCAGDTRSGGPQNLHLDACQGDSDGSLVCLNDGRMTLVGIISWGLGCGKDXVGYTYK 548
0Y 412 VSHFLPWTIRSHTK 424
Db 549 VTNVYLDWIRDNMR 561

```

RESULT 10
J05099
t-plasminogen activator (EC 3.4.21.66) beta precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999

Query Match	36.2%	Score 867.5;	DB 2,	Length 431;
Best Local Similarity	41.6%	Pred. No. 8,1e-59;		
Matches 185;	Conservative 61;	Mismatches 152;	Indels 47;	Gaps 10

```

QY 4 AARLLLVIVVSDK-----GSNELHQVPSNODCLANGCVSNKYFSNI 48
Db 9 LLVLLLCGAVFSLERQETRYOLANGSRAYGCSBLR-----CFNGTCMOAASFDF 61
QY 49 HWCNPKRFGGCHCEIDKSKTCYEGNNGHFYRKGASTIDTGRPCLPWNSATVLOQTYHAHR 108
Db 62 -VCGPKGYTGKQCEVDTHATCYKQGVYTRGTWSTSESAGACINNNSNLLTRRYNGRR 120
QY 109 SDALDLGKKNYGCNPNRRRPMKCYOVGLPLVCEWVHDCADKKRSPPEELKFOC 168
Db 121 SDATLTGIGNHNYCNPNDNSKPMCVIKASFLEIFECFPPVCS-----KATC 168
QY 169 G-QKTLRPRFKLIGSEFTTIEONPWFPAALYRRHRGGS-ITYVCGSGLISPCWVISAITHCF 226
Db 169 GLRKXKEPOLHSTGGLFTFDITSHPMQALFQNRNRSGBRFLCGGILISSCWLTAAHCF 228
QY 227 ID-YKKEDYIYVLRSLRNSVTQGBMKFEVENLLHKYASDTLAHNDIALLKTRKE 265
Db 229 QERYRPOHLRVV-LERTYRVKPKGKEQTFVEVKCIHIEFFDDT--YNNDIALLQLKSGS 285
QY 286 GRCAQPSRTIQTCLPSMYNDPQFGTSCETIFGKENSTDYLYPEOLKMTVVLLISHREC 345
Db 286 PQCAQESBSVRAICLPEANLQLPDWTBCELSGCKHKSXSPFYSBQLKGSHVRLYPSRSC 345
QY 346 QOPHYGSEVTTKMLCAADPOWKT-----DSCQGBSGRPLVCSLOGRNTLTGIVSKRG 399
Db 346 TSKPLFNFTVITNMMLCAGDTRSGEIIYPNVHDAQCGDSGSLVCMNDNMHTLLGIISWVG 405
QY 400 CALKDXKPGVYTVSHFLPWIRSHTK 424
Db 406 CGEKDIPGVYTKVTVYLGIMIRDNMR 430

```

RESULT 11
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (*Megaderma lyra*)
C.Species: *Megaderma lyra*
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #ext_change 10-Sep-1999
C.Accession: A34369
R.Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacobs
J. Biol. Chem. 264, 11947-11952, 1989
A.Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasminic
A.Reference number: A34369; MUID:90036867; PMID:2509450
A.Accession: A34369
A.Status: preliminary
A.Molecule type: mRNA

A:Residues: 1-477 <GAR>
A:Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis, glycoprotein, hydrolase, kringle, serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-47/Domain: plasminogen activator #status predicted <PA>
F:42-79/Domain: fibronectin type I repeat homology <IFA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KR>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.2%; Score 867.5; DB 1; Length 477;
Best Local Similarity 42.7%; Pred. No. 9,1e-59;
Matches 179; Conservative 61; Mismatches 150; Indels 29; Gaps 9;

QY 19 KGSNELHQVP-----SNCDLNGTCTVSNKYPFSNIMHCNCPKKGQHCEIDKSKTCYEGN 74
Db 74 RGLAQCHTVPYKSCSEIRCFNGTCTWQASFSDF-VCCPRKGYTGKCEVDTHATCYKDQ 132
QY 75 GHFPRGKASTDTMRPCLPANSAIVLQOTYHAHSDALQLGKXNTCRPNRRRRCWCY 134
Db 133 GVTYRGWTSSESGAQCINMNSNLTTRTYNGRSDAITLGLGNHNYCRPNDDNSKPCWCY 192
QY 135 VQVGLKPLVQECWHDCAADGKPPPELKFQCG-QKTLPRFKIIGGEFTTENOPIWF 193
Db 193 VYKSKITLERCSPVCS-----KATGRLKRYKPEQLHSHGGLFTDITSHPWQ 240
QY 194 AAIYRRHGGG-VTVYCGSLISPCWYISATHCPID-YPKKEDYIVYLGRSLNSNTQGE 251
Db 241 AAIYRRHGGG-VTVYCGSLISPCWYISATHCPID-YPKKEDYIVYLGRSLNSNTQGE 251
QY 252 MKFEVENLILHKYSAOTLAHNDIALIKIRSKGRCQAQSPRTQITCLPSMYNDPOPGT 311
Db 300 QTFEVEKCIYAEHEDDDT--YNDIALQLKSGSPQCAQESDSVRAICLPANQLPDMWT 357
QY 312 SCEITFGKENSSTYLYPEQLKMTVVKLISHRECOQPHYYGSEVTTKMLCAADPOMKT-- 369
Db 358 ECEISGCTGKHKSSPFISQLEKGVRLYPSRSTSKPLFKVITYTKMLCGDTRSGEIH 417
QY 370 ----DSCQDGGGLVCSLQGRMTLTGIVSWGRGKALDKDPVYTRYVSHFLPWIRSHTK 424
Db 418 PNVHADACQDGGGLVCSLQGRMTLTGIVSWGRGKALDKDPVYTRYVSHFLPWIRSHTK 476

RESULT 12

t-plasminogen activator (EC 3.4.21.68) precursor - rat
A:Residues: 1-559 <FEN>
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35029; A31597
R:Feing, P.; Ohlsson, M.; NY, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID:90130446; PMID:2105315
A:Accession: A35029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-559 <FEN>
A:Cross-references: GB:M31197; NID:g207429; PIDN:AAA42261.1; PID:g207431; GB:J05226
R:NY, T.; Leonardson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activat
A:Reference number: A31597; MUID:89170114; PMID:3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <NYT>
A:Cross-references: GB:M23667; NID:g530159; PIDN:AAA41812.1; PID:g530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis, glycoprotein, hydrolase, kringle, serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <IFI>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-394/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F:149,481/Binding site: carbohydrate (am) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Xile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 35.4%; Score 847.5; DB 1; Length 559;
Best Local Similarity 36.7%; Pred. No. 3,6e-57;
Matches 183; Conservative 65; Mismatches 153; Indels 97; Gaps 12;

QY 18 SKGSNELHQVP-----SNCDLNGTCTVSNKYPFSNIMHCNCPKKGQHCEIDKSKTCYEG 73
Db 69 NSGLAQCHSVYKSCSEIRCFNGTCTWQALYFSDF-VCCPRDFVGRCDIDTRATCFEG 127
QY 74 NGHPRGKASTDTMRPCLPANSAIVLQOTYHAHSDALQLGKXNTCRPNRRRRCWCY 133
Db 128 QGITYRGWTSSTANGAECINMNSNLTTRTYNGRSDAITLGLGNHNYCRPNDDNSKPCWCY 187
QY 134 VQVGLKPLVQECWHDCAADGKPPPELKFQCG-VHDCADGK-----KPSPP----- 161
Db 188 YVYKAKYITTEFCSTACPKPFTEDCYGKGYTRYRGTHSTTSKASCLPANSMLIGKTY 247
QY 162 -----EELKIQ-----CGQKTLR-PRF 177
Db 248 TAWRANSQALGLGRHNYCRPNDDAKPMWMDKRLKWEYCDMSPCSTGLRQYKOPQF 307
QY 178 KIIGGEFTTENOPIWPAALY-RHRRGSGTVYCGSLISPCWYISATHCPIDYPKKEDYI 236
Db 308 RIRGGFTDITSHPWQALIFVKNRSPGERFLCGVLISSCWLSAHCHEVERPPPHHLK 367
QY 237 VYGRSLNSNTQGEKMEFEVENLILHKYSAOTLAHNDIALIKIRSKGRCQAQSPRTQ 296
Db 368 VYGRSLNSNTQGEKMEFEVENLILHKYSAOTLAHNDIALIKIRSKGRCQAQSPRTQ 296
QY 297 TICLPSTMYNDPOF-----GTSCEITFGKENSSTYLYPEQLKMTVVKLISHRECOQPHYYG 352
Db 426 TACLP-----DPVQLPDMTECELSGYKHEASPPFSRDLKHAHVLVYPSRSTSKPLFKVITYTKMLCGDTRSGEIH 481
QY 353 SEVTTKMLCAADP-----QKTLPSRFKIIGGEFTTENOPIWF 193
Db 482 KTTTNNMLCGDTRTGNDVHADACQDGGGLVCSLQGRMTLTGIVSWGRGKALDKDPG 541
QY 408 VYTRYVSHFLPWIRSHTK 425
Db 542 IYTKVTRYVSHFLPWIRSHTK 559

RESULT 13

t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
J:050597
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: J050597
R:Krauszschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; Dom
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: J050597; MUID:92039036; PMID:1937019
A:Accession: J050597
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis, glycoprotein, hydrolase, kringle, serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>

F/22-36/Domain: propeptide #status predicted <PRO>
 F/37-47/Product: plasminogen activator alpha-1 #status predicted <PLA>
 F/42-79/Domain: fibronectin type I repeat homology <1RA>
 F/87-120/Domain: kring homology <EGF>
 F/128-209/Domain: kring homology <KRG>
 F/226-471/Domain: trypsin homology <TRY>
 F/42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-245,257-273,265-334,359-4
 F/153,398/Binding site: carbonyl dipeptide (Asn) (covalent) #status predicted
 F/225-226/Cleavage site: His-Ser (plasmin) #status predicted
 F/272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 35.2%; Score 841.5; DB 2; Length 477;
 Best Local Similarity 41.8%; Pred. No. 8.8e-57;
 Matches 176; Conservative 57; Mismatches 151; Indels 37; Gaps 10;

19 KGSNELHQVPSN-CD---CLNGGTCVSNKYPFSNHNHCNCPKKGQHCIDKSTCYEGN 74
 74 RQGARCHTVVNSCSSEPCFNGGTCWQAVYFSDF-VCCPCGYTGKCEVDTRATCYEQ 132
 75 GHFPGKASTDTMGRPCLPMNSATVLCQTYHAHRSALQGLGKHNCRPNDRRPMCY 134
 133 GVTYRGTWSTAESRVECEINNNSSILTRITNGRMDAFNIGLGNHNYCRPNGAPKPMCY 192
 135 VQVGLKPLVQECMVHDCADGKKSPPEELKPOCG-OXTRPRPKXIIGSEFTTENQPMF 193
 193 VIKAGKFTSSECSVPVCS-----KATGRLKRYKEPOLHSTGGLFTDITSHPMQ 240
 194 AAIYRRHRRGS-VTVVCGGSLSPCWYISATHGCIIDPKKEDI-----VILGRSLNSN 247
 241 AAIYRRHRRGS-VTVVCGGSLSPCWYISATHGCIIDPKKEDI-----VILGRSLNSN 247
 248 TQGEKMEFEVNLILKDYASDTLAHNDIALKIRSKRCAQPSRTITCLPSMTYNDP 307
 296 GEEGQFKVKKYIVHKEFDDT--YNNDIALLQKSPCCQAQESDSVRAICLPEALQL 353
 308 QGFTSCETGFGKENSVDLYLPBOLKMTVVKLISHRECOQPHYYSFTTKMLCAADPQW 367
 354 PDWTECELSGKSSPFYSEQLKEGHVRLYPSRCAPKFLFNKVTNNMLCAGDTR 413
 368 KT-----DSGCGDSGGLVCSLQGMRTLTGIVSMGCGALCKDKPGYTVSHFLPIRS 421
 414 GEIYPMVHDCQSDSGGLVCMNDHMTLTGLIISWVGCGEKDVPGYTVTVYLGWIRD 473
 422 H 422
 474 N 474

RESULT 14

t-plasminogen activator (EC 3.4.21.68) precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A29941; S48205; S48206
 R/Rickles, R.J.; Darrow, A.L.; Strickland, S.
 J. Biol. Chem. 263, 1563-1569, 1988
 A/Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator m
 A/Reference number: A29941; MIMD:86087303; PMID:2826484
 A/Accession: A29941

A/Molecule type: mRNA
 A/Residues: 1-559 <RIC>
 A/Cross-references: GB:J03520; NID:G202109; PIDN:AAA40470.1; PID:G202110
 Eur. J. Biochem. 224, 863-871, 1994
 A/Title: Characterization of the murine plasma fibrinolytic system.
 A/Reference number: S48202; MIMD:95010076; PMID:7523120
 A/Accession: S48205
 A/Molecule type: protein
 A/Residues: 33-37,'X',39-40 <LIU>
 A/Accession: S48207
 A/Molecule type: protein
 A/Residues: 309-316 <LIU>
 A/Accession: S48206

A/Molecule type: protein
 A/Residues: 33-37,'X',39-40 <LIU>
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 A/Keywords: fibrinolysis; glycoprotein; hydrolase; kring; serine proteinase
 F/1-17/Domain: signal sequence #status predicted <SIG>
 F/18-29/Domain: propeptide #status predicted <PRO>
 F/30-559/Product: t-plasminogen activator #status predicted <MAT>
 F/330-368/Product: t-plasminogen activator chain A #status predicted <ACH>
 F/149,481/Binding site: carbonyl dipeptide (Asn) (covalent) #status predicted
 F/308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
 F/335,404,510/Active site: His, Asp, Ser #status predicted

Query Match 34.8%; Score 833.5; DB 1; Length 559;
 Best Local Similarity 35.7%; Pred. No. 4.3e-56;
 Matches 178; Conservative 66; Mismatches 157; Indels 97; Gaps 10;

18 SKSNEIHOVP---SNDCNGTCVSNKYPFSNHNHCNCPKKGQHCIDKSTCYEG 73
 69 NSGLVQCHSVPRVSCSEPCFNGGTCQALYFSDF-VCCPCGYTGKCEVDTRATCYEQ 127
 74 NGHFGKASTDTMGRPCLPMNSATVLCQTYHAHRSALQGLGKHNCRPNDRRPMCY 133
 128 QGITYRGTWSTAESRVECEINNNSSVILKRYNARRPVALKGLGNHNYCRPNDRPKM 187
 134 YVQVGLKPLVQECMVHDCADGKKSP-----FQCG-OXTRPRF 158
 188 YVRKAKYTFEPCSTACPRKSEDCVGGVYTRGTHSLTTSQASCLPMNSIVLMGSKY 247
 159 -----SPPEELK-----FQCG-OXTRPRF 177
 248 TAWRTNSQALGLARHNYCRPNDDARPMCKDKRLTWECDSMPCSTGLRQYKRPQF 307
 178 KIIGSEFTTENQPMFPAIY-RHRRGSVTVVCGGSLSPCWYISATHGCIIDPKKEDI 236
 308 RINGGLYTDITSPWQALFVKNKRSPPGERFLCGVLISSCWLSAACHLEPPNNHLK 367
 237 VILGRSLNSNTQGEKMEFEVNLILKDYASDTLAHNDIALKIRSKRCAQPSRTITQ 236
 368 VILGRYRVVPGSEEGQTFEIKYIVHEFPDDT--YNDIALLQKSPCCQAQESDSV 425
 297 TILPSMYNDPQF---GTSCEITGFGKENSVDLYLPBOLKMTVVKLISHRECOQPHY 352
 426 TACLTP---DPNQLPDWTECELSGKHSAPFSRDLKEAHVRLYPSRCTSQHFLN 481
 353 SEVTTMGLCAADP-----QMKTDSGCGDSGGLVCSLQGMRTLTGIVSMGCGALCK 407
 482 KVTNNMLCAGDTRSGNDLHDACQSDSGGLVCMINKMTLTGLIISWGLGCGQKDPV 541
 408 VYTVSHFLPWSHTKE 425
 542 VYTKVTNYIDMIDNMKQ 559

RESULT 15

t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
 C/Superfamily: tissue plasminogen activator
 N/Alternate names: tissue plasminogen activator
 C/Species: Desmodus rotundus (common vampire bat)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 R/Kraatzschmar, J.; Haendler, B.; Langer, G.; Böldol, W.; Bringmann, P.; Alagon, A.; Domr
 Gene 105, 229-237, 1991
 A/Title: The plasminogen activator family from the salivary gland of the vampire bat Desm
 A/Reference number: J050597; MIMD:92039036; PMID:1937019
 A/Accession: J050600
 A/Molecule type: mRNA

A:Residues: 1-394 <KRA>
A:Cross-references: GB:M63990; NID:G166078; PIDN:AAA31595.1; PID:G166079
A>Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:122-36/Domain: propeptide #status predicted <PRO>
F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
F:143-368/Domain: kringle homology <KRG>
F:45-126/Domain: trypsin homology
F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189,238,345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.4%; Score 752; DB 2; Length 394;
Best local similarity 42.0%; Pred. No. 5,1e-50;
Matches 155; Conservative 56; Mismatches 134; Indels 24; Gaps 7;
QY 65 DKSRTCEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGKHNKCN 124
Db 40 DPHATCYDQGVTRGTWTSTESGAQCIWNNSNLIRTYNGRMPEAVKLGHNKCN 99
QY 125 PDNRPRPCTVQVGLKPLVQECWVHDCADGKKSSPPEELKFCG-QKTRPRFKITIGE 183
Db 100 PDGASKPWCYVYKARKFTSESCSVPCS-----KATGALKRYKEPOLHSTGGL 147
QY 184 FTTIENOPFAIYRRHRGS-VTVVCGSLISPCWVISAATHCFID-YPKKEDYIVYLGR 241
Db 148 FTDITSHPMOAIYRQNRSSGERFLCGILISSCWVLTAAHCFQERYPPQHLRV-LGR 206
QY 242 SRLNSNTQGENKFEVENLIHKDYSADTLAHNDIALIKIRSEKRGCAQPSRTIQTICLP 301
Db 207 TYRVKPGKEQTFVEKCIHVEEFDPT--YNNIDIALQLKSGSPQCAQESDSVRAICLP 264
QY 302 SMVNDPQFGTCETITGPKENSTDYLPQULKMTVVKLISHREGQPHYIGSEVTTMLC 361
Db 265 EANTQLPDMTECELSGYGKHSSPFYSEQLKEGHVRLYPSSRCTSKFLFNKTYTNMMLC 324
QY 362 AADPOMKT-----DSCQDSGCPVYCSLQGRMTLTGIVSGRGCALKDXKGVYTRVSHF 415
Db 325 AGDTRSGEIIYPNVHDAQCQDSGGFLVCMNDNMHTLLGIISWVGCGEKDIPGVYTKVINY 384
QY 416 LPWIRSHTK 424
Db 385 LGWIRDNMNR 393

Search completed: March 18, 2004, 11:20:16
Job time : 45.326 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 10:54:45 ; Search time 26.407 Seconds

(without alignments)
849,859 Million cell updates/sec

Title: US-10-076-421-2

Perfect score: 2394
Sequence: 1 MRALLRLLLCVLVSDSKG.....VSHFLPWRSHKENGALL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2391	99.9	431	1	UROK_HUMAN
2	2228	93.1	433	1	UROK_PAPCY
3	1928.5	80.6	442	1	UROK_PIG
4	1835	76.6	433	1	UROK_BOVIN
5	1728.5	72.2	432	1	UROK_RAT
6	1700	71.0	433	1	UROK_MOUSE
7	1009	42.1	434	1	UROK_CHICK
8	868.5	36.3	477	1	URT2_DESRO
9	868.5	36.3	562	1	TPA_HUMAN
10	867.5	36.2	431	1	URT2_DESRO
11	847.5	35.4	559	1	TPA_RAT
12	841.5	35.2	477	1	URT1_DESRO
13	828.5	34.6	559	1	TPA_MOUSE
14	821	34.3	566	1	TPA_BOVIN
15	752	31.4	394	1	URT2_DESRO
16	735.5	30.7	654	1	HGFA_HUMAN
17	727	30.4	653	1	HGFA_MOUSE
18	718.5	30.0	603	1	FA12_MOUSE
19	692	28.9	615	1	FA12_HUMAN
20	643.5	26.1	533	1	FA12_HUMAN
21	504.5	21.1	790	1	PLMN_PIG
22	501	20.8	812	1	PLMN_MOUSE
23	499	20.8	810	1	PLMN_HUMAN
24	497	20.8	810	1	PLMN_MOUSE
25	494.5	20.7	833	1	PLMN_MOUSE
26	493	20.6	833	1	PLMN_MOUSE
27	484.5	20.2	812	1	PLMN_MOUSE
28	474	19.8	1420	1	APOA_MOUSE
29	474	19.8	1420	1	APOA_MOUSE
30	462	19.3	875	1	NETR_HUMAN
31	449.5	18.8	338	1	PLMN_MOUSE
32	445	18.6	761	1	NETR_MOUSE
33	435	18.2	810	1	PLMN_MOUSE

34	424.5	17.7	418	1	HATT_HUMAN
35	420	17.5	811	1	TMS6_HUMAN
36	419.5	17.5	811	1	TMS6_MOUSE
37	417	17.4	436	1	HEPS_MOUSE
38	410	17.1	436	1	KAL_MOUSE
39	408.5	17.1	343	1	PSS6_HUMAN
40	406	17.0	711	1	HGFL_HUMAN
41	405.5	16.9	855	1	ST14_HUMAN
42	405.5	16.9	855	1	ST14_MOUSE
43	405	16.9	277	1	KLKD_HUMAN
44	404.5	16.9	455	1	TMS5_MOUSE
45	402.5	16.8	417	1	HEPS_HUMAN

ALIGNMENTS

RESULT 1
UROK_HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q16618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
GN (U-plasminogen activator).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=85215647; PubMed=2987867;
RX Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP MEDLINE=85215647; PubMed=2987867;
RX Holmes W.E., Pennica D., Blaber M., Roy M.W., Guenzler W.A.,
RA Steffens G.J., Heynaker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP MEDLINE=86056954; PubMed=2415429;
RX Nagai M., Hiratake T., Kaneda T., Hayasuke N., Arimura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
RL Gene 36:183-188(1985).
RN [4]
RP MEDLINE=86056954; PubMed=2415429;
RX Jacob P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";
RL DNA 4:139-146(1985).
RN [5]
RP MEDLINE=85203359; PubMed=3888571;
RX Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Peol C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP MEDLINE=22388257; PubMed=12477932;
RX Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Stransberg R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang D., Heien F.,
Diatchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carmini P., Prange C.,
 RA Bawa S.S., Loguano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Keltman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schermer A., Schein J.B., Jones S.J.M., Maria M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [17]
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.,
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RT Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [18]
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Oetting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.,
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RT Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [19]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
 RA Studer R.O.,
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RT Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Oetting F., Frankus E., Flohe L.,
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RT Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1056(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones B.Y.,
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RT Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Speil S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.,
 RT "4-aminomethylphenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RT Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberg M., Smith R.A.G., Dobson C.M.,
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RT Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.,

RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RT Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=9419701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas W., Smith R.A.G., Dobson C.M.,
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RT J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sawaaki Y., Hanada K.,
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure.";
 RT Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Conne B., Berczy M., Belin D.,
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RT Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Conne B., Berczy M., Belin D.,
 RT Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.,
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RT Electrophoresis 18:686-689(1997).
 RN [20]
 RP FUNCTION: Potent plasminogen activator and is clinically used for
 CC therapy of thrombolytic disorders.
 CC
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain.
 CC
 CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in pulmonary embolism (PE) to initiate fibrinolysis.
 CC
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X02419; CAA26268.1; -
 CC EMBL; M15476; AAA61253.1; -
 CC EMBL; D00244; BAA00175.1; -
 CC EMBL; D11143; BAA01919.1; -
 CC EMBL; X02760; CAA26535.1; -
 CC EMBL; AF377330; AAK53822.1; -
 CC EMBL; BC013575; AA113575.1; -
 CC EMBL; X03226; AAC97138.1; -
 CC EMBL; K03286; AAA61252.1; -
 CC EMBL; A21571; CAA01559.1; -
 CC EMBL; A18397; CAA01390.1; -
 CC PIR; A00931; UKHU.
 CC PDB; 1KDU; 31-OCT-93.

Query Match 99.9%; Score 2391; DB 1; Length 431;
 Best Local Similarity 99.8%; Pred. No. 6.6e-182;
 Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLLARLLLCVIVSDSKSNELHGVPSNCDLNGTCVSNKYFSNIHNCNPKKFGGQ 60
 DB 1 MRLLARLLLCVIVSDSKSNELHGVPSNCDLNGTCVSNKYFSNIHNCNPKKFGGQ 60
 QY HCEIDSKTCYBENGHFRGKASTDTMGRPCLPWNASTVYQOTYHARSALOLGJGKN 120
 DB HCEIDSKTCYBENGHFRGKASTDTMGRPCLPWNASTVYQOTYHARSALOLGJGKN 120
 QY 121 YCNPNRRRPMWCYVVGKPLVQECMVHDCADGKKSPPEELKFCGCKTLRPRKII 180
 DB 121 YCNPNRRRPMWCYVVGKPLVQECMVHDCADGKKSPPEELKFCGCKTLRPRKII 180
 QY 181 GGEFTTIENQPMFALYRRHGGSVTVYCGGSLISPCWVISATHCFTIDYKEDYIYVLG 240
 DB 181 GGEFTTIENQPMFALYRRHGGSVTVYCGGSLISPCWVISATHCFTIDYKEDYIYVLG 240
 QY 241 RSLNNTQSGEMKEVENLLHMDYADTLAHNDIALTKIRSEGCAPSRPTOTICL 300
 DB 241 RSLNNTQSGEMKEVENLLHMDYADTLAHNDIALTKIRSEGCAPSRPTOTICL 300
 QY 301 PSWYNDPOFGTSCETIFGKENDTDLYPEQLKMTVVKLISHRECOQPHYGSEVTKML 360
 DB 301 PSWYNDPOFGTSCETIFGKENDTDLYPEQLKMTVVKLISHRECOQPHYGSEVTKML 360
 QY 361 CAADPMKTDSCQDGGSGPLVCSLQGMVLTGIVSWRGALADKDPGVTVRVSHFLPWIR 420
 DB 361 CAADPMKTDSCQDGGSGPLVCSLQGMVLTGIVSWRGALADKDPGVTVRVSHFLPWIR 420
 QY 421 SHTKEENGLAL 431
 DB 421 SHTKEENGLAL 431

RESULT 2
 ID UROK_PAPCY STANDARD; PRT; 433 AA.
 AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (BC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Papio.
 OX NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RA MEDLINE=90287734; PubMed=2113276;
 RA Au Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RT plasminogen activator."
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain (by similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 CC EMBL; X51935; CAA6200.1; -.
 CC PIR; S14687; URBAY.
 CC HSSP; P00749; 1LMW.
 CC MEROPS; S01.231; -.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR008293; Pept_S1A_uPA.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00051; Kringle1; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PIRSF; PIRSF001144; Urk_plasm_act; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; TRYP_SPC; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; FALSE_NEG.
 CC PROSITE; PS00026; EGF_3; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00240; KRINGLE_2; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC K1M Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
 CC K1M Kringle; EGF-like domain; Zymogen; Signal.
 CC FT SIGNAL 1 20
 CC FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 CC FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
 CC FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 CC FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 CC FT DOMAIN 26 62 EGF-LIKE.
 CC FT DOMAIN 69 150 KRINGLE.
 CC FT DOMAIN 151 177 CONNECTING PEPTIDE.
 CC FT DOMAIN 178 433 SERINE PROTEASE.
 CC FT DISULFID 30 38 BY SIMILARITY.
 CC FT DISULFID 32 50 BY SIMILARITY.
 CC FT DISULFID 52 61 BY SIMILARITY.
 CC FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 CC FT DISULFID 208 224 BY SIMILARITY.
 CC FT DISULFID 216 287 BY SIMILARITY.
 CC FT DISULFID 315 384 BY SIMILARITY.
 CC FT DISULFID 347 402 BY SIMILARITY.
 CC FT ACT SITE 223 223 CHARGE RELAY SYSTEM.
 CC FT ACT SITE 224 274 CHARGE RELAY SYSTEM.
 CC FT ACT SITE 378 378 CHARGE RELAY SYSTEM.
 CC FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
 CC SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 93.1%; Score 2228; DB 1; Length 433;
 Best Local Similarity 92.4%; Pred. No. 5.2e-169;
 Matches 401; Conservative 17; Mismatches 12; Indels 4; Gaps 2;

QY 1 MRLLARLLLCVIVSDSKSNELHGVPSNCDLNGTCVSNKYFSNIHNCNPKKFGGQ 60
 DB 1 MRLLARLLLCVIVSDSKSNELHGVPSNCDLNGTCVSNKYFSNIHNCNPKKFGGQ 59
 QY HCEIDSKTCYBENGHFRGKASTDTMGRPCLPWNASTVYQOTYHARSALOLGJGKN 120
 DB HCEIDSKTCYBENGHFRGKASTDTMGRPCLPWNASTVYQOTYHARSALOLGJGKN 119
 QY 121 YCNPNRRRPMWCYVVGKPLVQECMVHDCADGKKSPPEELKFCGCKTLRPRKII 180

```

Db 120 YCENPDNRPRPCVYVGLKQVOCVHNCAGKSSPPEEIOFCGQRTLRPRKIV 179
QY 181 GGEFTTENPWPFAIYRRHRGSGVTVVCGSLISPCWVISATMCPFDYKKEIDYIYL 240
Db 180 GGEFTTENPWPFAIYRRHRGSGVTVVCGSLISPCWVISATMCPFDYKKEIDYIYL 239
QY 241 RSLNSTQGEKMEKEVENTLHKDYSADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
Db 240 RSLNSTQGEKMEKEVENTLHKDYSADTLAHNDIALKIRSKERCAQPSRTIQTICL 299
QY 301 PSWYNDPQ---FGTSCITGFGKENSVDYLYPEOLKMTVVKLISHRECOOPHYGSEVTT 357
Db 300 PSWYNDPNDPFPFSGCEITGFGKENSVDYLYPEOLKMTVVKLISHRECOOPHYGSEVTT 359
QY 358 KMLCAADPOKMTSCQGDSDGSLVCSIQGEMTLTGIVSWRGKALKDKPGVYTRVSHFLP 417
Db 360 KMLCAADPOKMTSCQGDSDGSLVCSIQGEMTLTGIVSWRGKALKDKPGVYTRVSHFLP 419
QY 418 WIRSHTEKENGAL 431
Db 420 WIRSHTEKENGAL 433

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Negamine Y.; Pearson D.; Altus M.S.; Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Negamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL; X01648; CA25806.1; -.
DR EMBL; X02724; CA26511.1; -.
DR PIR; A00932; UKPG.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR InterPro; IPR001254; Peptidase_S1.

```

```

DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PIRSF; PIRSF001144; Uro_kinase_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR_1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50020; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW kringle; EGF-like domain; Zymogen; Signal.
KW
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 299
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 387 387
FT ACT_SITE 241 241
FT CONFLICT 242 242
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; EE32FCFE50132EE CRC64;

Query Match 80.64; Score 1928.5; DB 1; Length 442;
Best Local Similarity 79.4%; Pred. No. 2,7e-145; Indels 11; Gaps 2;
Matches 350; Conservative 33; Mismatches 47;

QY 1 MRALLRLALCVLVSDSKSGSNELHGV--PSNDCINGGTGVSNKYFNSIHWNCNCPKFG 58
Db 1 MRLVLRACLSCVLVSDSKSGSHELHQSASNGCGLNGKCVGKYFNSIQRCSCKPKFG 60
QY 59 GQCEIDKSKTCYEGNGHFPKRASTDTYMGKRPCLPNNASATVLOOTYHAHSDALQGLGK 118
Db 61 GEHCEIDTSQTCCEGNHSGYKRAANTYGRPCLPNNASATVLTNTYHAHSDALQGLGK 120
QY 119 HNYCRNDNRPRPCVYVGLKQVOCVHNCAGKSSPPEEIOFCGQRTLRPRKIV 169
Db 121 HNYCRNDNRPRPCVYVGLKQVOCVHNCAGKSSPPEEIOFCGQRTLRPRKIV 180
QY 170 QKTIRPRFKIIGSEFTTENPWPFAIYRRHRGSGVTVVCGSLISPCWVISATMCPFDY 229
Db 181 QKALRPRFKIIGSEFTTENPWPFAIYRRHRGSGVTVVCGSLISPCWVISATMCPFDY 240
QY 230 PKKEDYTVYIGSRSLNSTQGEKMEKEVENTLHKDYSADTLAHNDIALKIRSKERCA 289
Db 241 QOKEDYTVYIGSRSLNSTQGEKMEKEVENTLHKDYSADTLAHNDIALKIRSKERCA 300
QY 290 QPSRTIQTICLPSWYNDPQSGCEITGFGKENSVDYLYPEOLKMTVVKLISHRECOOPH 349
Db 301 QPSRSIQTIQLPVPNGAHFGASCEIVGFKEDPSDYLYPEOLKMTVVKLISHRECOOPH 360

```

QY 350 YVESVTTKMLCAADPQMTKDSGCGSGGPLYCSLQGRMTITGVISWGRGALKDKPKGVY 409
 DB 361 YVESVTTKMLCAADPQMTKDSGCGSGGPLYCSLQGRMTITGVISWGRGALKDKPKGVY 420
 QY 410 TRVSHFLPWIRSHKENGSLA 430
 DB 421 TRVSRFLTWIHTHVGENGSLA 441

RESULT 4
 UROK_BOVIN STANDARD; PRT; 433 AA.

AC 005359; Q28209;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 GN (U-plasminogen activator).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovine; Bos.
 NC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=93216119; PubMed=8385052;
 RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
 RA Schlunberg M.-D.;
 RT "Bovine urokinase-type plasminogen activator and its receptor:
 RT cloning and induction by retinoic acid.";
 RL Gene 125:177-183(1993).
 RN [2]

RP SEQUENCE OF 12-433 FROM N.A.
 RC TISSUE=Kidney;
 RA Ravn P., Berglund L., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen activators uPA
 RT and tPA.";
 RL Int. Dairy J. 5:605-617(1995).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- INDUCTION: By retinoic acid.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: L03546; AAA51419.1; -
 DR EMBL: X85801; CAAS9796.1; -
 DR PIR: JN0560; JN0560.
 DR HSSP: P00749; ILMW.
 DR MEROPS: S01.231; -
 DR InterPro: IPR009003; Cys_ser_crypshn.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR008293; Pept_S1A_uPA.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00089; trypsin; 1.
 DR Pfam: PF00051; kringle; 1.
 DR PIRSF: PIRSF001144; uPA plasm act; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.

DR SMART; SM00020; TRYP_SPE; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle, EGF-like domain; signal; zymogen.

FT SIGNAL 1
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
 FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 180 CONNECTING PEPTIDE.
 FT DOMAIN 181 433 SERINE PROTEASE.
 FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 290 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 189 189 A -> T (IN REF. 2).
 SQ SEQUENCE 433 AA; 48730 MW; 4DE1BBD4DA47027A CRC64;

Query Match 76.6%; Score 1835; DB 1; Length 433;
 Best Local Similarity 75.1%; Pred. No. 6; ee-138;
 Matches 325; Conservative 46; Mismatches 60; Indels 2; Gaps 1;

QY 1 MRALLRLILCVAVSPDSKSNELHGV--PSNCDCLNGCGCVSNKYSNTIHWNCNCPKFG 58
 DB 1 MRVLLACLVALVSDSDSNEVHKGESGNCGLNGKCVYKYSNIQRSCPKFQ 60

QY 59 GORCEIDSKTCYEGNGHFYRGASTDTWGRPCLPMNSATVLOQTYAHRSDALQGLGK 118
 DB 61 GEHCIEDTSKTCYQNGHSGYRGKANDLSGRPCLANDSPVLLKMYAHRSDAIQLGLGK 120

QY 119 HNYCRNPNRRRWCVYQVGLKPLVQECVHDCADGKSSPPEBELKFCGCKTLRRFK 178
 DB 121 HNYCRNPNDRRWCVYQVGLKPLVQECVHDCADGKSSPPEBELKFCGCKTLRRFK 180

QY 179 IIGGEFTTENOPWPAIYBRHGGSTVYCGSLISPCVVISATNHCFTIDYPKKEDYIVY 238
 DB 181 IYGGQVTNANQWPAIYBRHGGSTVYCGSLISPCVVISATNHCFTIDYPKKEDYIVY 240

QY 239 IGRSRLNSNTGEMKEVENTILHKDYADTLAHNDIALKLRSGKRCAPSPRTIQT 298
 DB 241 LGSRSLNSNTGEMKEVENTILHKDYADTLAHNDIALKLRSGKRCAPSPRTIQT 300

QY 299 CLPSPMTNDPOFGTSCITGFGKENSVDLYPEOLKNTVVKLISHRECQPHYIGSEVTTK 358
 DB 301 CLPPEHEDHARSRTCEITGFGKENSVDLYPEOLKNTVVKLISHRECQPHYIGSEVTTK 360

QY 359 MLCADPQMTKDSGCGSGGLVCSLQGRMTITGVISWGRGALKDKPKGVYTVSHFLPW 418
 DB 361 MLCADPQMTKDSGCGSGGLVCSLQGRMTITGVISWGRGALKDKPKGVYTVSHFLPW 420

QY 419 IRSHTEKENGSLAL 431
 DB 421 INTHTGELNVLV 433

RESULT 5
 UROK_RAT

ID UROK_RAT STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN (U-plasminogen activator).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
RA Kefford R.F.,
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RL Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 156 in the low
CC molecular mass form to yield a short A1 chain (By similarity).
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63434; CAA45028.1; -;
CC EMBL; X65651; CAA46601.1; -;
CC PIR; S24604; S18932.
CC HSSP; P00749; IKDU.
DR MEROPS; S01.231; -;
DR InterPro; IPR000903; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19
FT CHAIN 20 432
FT CHAIN 20 177
FT CHAIN 156 177
FT CHAIN 179 432
FT CHAIN 27 63
FT DOMAIN 70 151
FT DOMAIN 152 178
FT DOMAIN 179 432
FT DISULFID 31 39
FT DISULFID 33 51
FT DISULFID 53 62
FT DISULFID 168 300
FT DISULFID 210 226
FT DISULFID 218 289
FT DISULFID 314 383
FT DISULFID 346 362
FT DISULFID 373 401
FT ACT_SITE 225 225
FT ACT_SITE 276 276
FT ACT_SITE 377 377
FT ACT_SITE 16 16
FT CONFLICT 24 24
FT CONFLICT 332 332
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;
Query Match 72.2%; Score 1728.5; DB 1; Length 432;
Best Local Similarity 71.3%; Pred.No.1,8e-129;
Matches 308; Conservative 51; Mismatches 70; Indels 3; Gaps 3;

QY 1 MRALLRLLLCVLVVSDSKSNEH-QVPSNCDCLNGKGVSNKYSNTHMCMCPKPKFG 59
1 MRWLASLPLCAL-VANSEGSLEASDSNCGCGVGVSVYKYSIRRCSPKPKFG 59
DB 60 QHCEIDSKTCYHGNGHFFRGKASTDTWGRPCLIWNASATVLQOTYHAHSDALQLGLGRH 119
60 EHCEIDTSKTCYHGNGQSVRGKANTDTKGRPCLIWNASPAVLQOTYNAHRSDALSLGLGRH 119
QY 120 NTCRNDRNRPRPCYOVGLKPLVORCM/HDGDKKSPPEBLKFCGCGKTLRRFRKI 179
120 NTCRNDRNRPRPCYOVGLKQPFVORCMVODSLSKSPSYDQGFQCGQKLRFRFKI 179
DB 180 IGGEFTTIENQPFALTYRRHGRGS-VTVYCGSLISPCWVISATHCFIDYPKEDYIVY 238
180 VGGEFTTVENQPFALTYLKNKGSPSPKCGSLISPCWVASATHCFVNPKEKEYIVY 239
QY 239 LGRSLNSNTQGMKEVENLIHKQYADTLAHNDIALLKIRSKRGCAOPSRTIOTT 298
240 LGSKRNQSYNPGEMKEVEQLILHEDFSDETLAFHNDIALLKIRISTGCAOPSRTIOTT 299
DB 299 CLPSWYNDQFGFSCITGFGKNSDYLYPBLKMTVVKLISHRECOOPHYVGSVYTK 358
300 CLPFRFGDAPFSDCITGFGQESATDYFPYKOLKNSVVKLISHRECKOPHYVGSINYK 359
QY 359 MLCADPQWKTDSCQDSCGPIVCSLQGRMTLGLVSMRGCALXDKPQVYTVRVSHFLPW 418
360 MLCADPEWKTDSCSDSGPILCNIDGRPTLSGLVSMGSCALXKXKPPVYTVRVSHFLPW 419
DB 419 TNSHTKEENGCLA 430
420 IQSHIGEENGCLA 431
DB

RESULT 6
UROK_MOUSE
ID UROK_MOUSE STANDARD; PRT; 433 AA.
AC P06869;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

(U-plasminogen activator).
 GN PLAU.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=65179474; Pubmed=2985383;
 RA Belin D., Vassalli J.-D., Combeppone C., Godeau F., Naganine Y.,
 RA Reich E., Kocher H.P., Duvolsin R.M.,
 RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
 RT mouse urokinase-type plasminogen activator.";
 RL Eur. J. Biochem. 148:225-232(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88163489; Pubmed=2831940;
 RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.,
 RT "The murine urokinase-type plasminogen activator gene.";
 RL Biochemistry 26:8270-8279(1987).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X02389; CAA26231.1; -
 CC EMBL: M17923; AAA40539.1; -
 DR PIR: A29420; UKMS.
 DR HSSP: P00749; IKDU.
 DR MEROPS: S01.231; -.
 DR MGD: MG1:97611; Plau.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR008293; Pept_S1A_uPA.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF000051; kringle_1.
 DR Pfam: PF000089; trypsin_1.
 DR FIRST: FIRST001144; Urk_plaem_act; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle_1.
 DR SMART: SM00181; EGF_1.
 DR SMART: SM00130; KR_1.
 DR SMART: SM00020; TYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; FALSE_NEG.
 DR PROSITE: PS50026; EGF_3; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS50070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KM Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KM kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 178 CHAIN A (BY SIMILARITY).

FT CHAIN 157 178 SHORT A CHAIN (A1).
 FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 28 64 EGF-LIKE.
 FT DOMAIN 71 152 KRINGLE.
 FT DOMAIN 153 179 CONNECTING PEPTIDE.
 FT DOMAIN 180 433 SERINE PROTEASE.
 FT DISULFID 32 40 BY SIMILARITY.
 FT DISULFID 34 52 BY SIMILARITY.
 FT DISULFID 54 63 BY SIMILARITY.
 FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 290 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 402 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
 FT ACT_SITE 227 227 CHARGE RELAY SYSTEM.
 FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
 FT SEQUENCE 433 AA; 48268 MW; A99C35F625043F9 CRC64;
 SQ
 Query Match 71.0%; Score 1700; DB 1; Length 433;
 Best Local Similarity 69.2%; Pred. No. 3 2e-127; Indels 2; Gaps 2;
 Matches 299; Conservative 56; Mismatches 75;
 QY 1 MRLLARLLLCVLVSDSKGSENLHGV-PSNCDLNGTCVSNKRYSEHMCNCPKKEGG 59
 DB 1 MKWVLASFLFCALVVKNSBEGSVILGAPDESNCCQNGVCVSKYFSRIRRCGCPKRFQ 60
 QY 60 QHCEIDSKTCYSGNHFYRGRKASTDTMGAPPCIPNNSATVLOOTYAHREBDALQIGK 119
 DB 61 EHCIEDASKTCYGNDSYRGKANTDTKGRPCCLAMNAPALQRYNARPDALSLGK 120
 QY 120 NTCRNPDMRRPCYQVGLKPLVQECMYDCADGKKSPPELKPQCGKTLRPFK 179
 DB 121 NTCRNPDMRRPCYQVGLKPLVQECMYDCADGKKSPPELKPQCGKTLRPFK 180
 QY 180 IGEFTTIENQPFAPAIYRRHRGS-VTVYCGSLISPCWVISATHCFIDPKKEDVIY 238
 DB 181 VGEFTTEVENQPFAPAIYQNKKGSPSPFCGSLISPCWVAALCFIDPKKENVIV 240
 QY 239 LGRSLNSNTQSGMKREVENLIHKDYADTLAHNDIALKTRSKRGCAQPSRTQT 298
 DB 241 LQSGKSSSYNPGEMKEVEBOLILHEYYREDSLAYHNDIALKTRTSYGCAQPSRSYOT 300
 QY 299 CLPSMNDPQFGNSCRITGKENSSTDYLYPEQLKMTVVYLISHRQOQPHYGSFVTK 358
 DB 301 CLPFRFTDAPFGSDCEITGKESBSDLYPKLAKMSVVLVSHBQOMQPHYIGSELINX 360
 QY 359 MLCADPQWKTDSCQDSSGGLVCSLQGRMTLTGIYSWGRGALKDKPGVYTRVSHLPW 418
 DB 361 MLCADPQWKTDSCQDSSGGLVCSLQGRMTLTGIYSWGRGALKDKPGVYTRVSHLPW 420
 QY 419 TRSHTEENGCLA 430
 DB 421 IOSHIGEGKCLA 432
 RESULT 7
 UROK_CHICK STANDARD; PRT; 434 AA.
 AC P5120;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.

CC MEDLINE=90110185; PubMed=2295632;
 RA Leelle N.D., Kessler C.A., Bell S.M., Degen J.L.;
 RT "The chicken urokinase-type plasminogen activator gene";
 RL J. Biol. Chem. 265:1339-1344(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; J05187; AAA49131.1; -;
 DR EMBL; J05188; AAA49130.1; -;
 DR PIR; A35005; A35005.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.231.1.1;
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_1like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle_1.
 DR Pfam; PF00089; trypsin_1.
 DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle_1.
 DR SMART; SM00181; EGF_1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Src; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Zymogen.
 FT SIGNAL 1 20
 FT CHAIN 21 434
 FT CHAIN 21 171
 FT CHAIN 173 434
 FT DOMAIN 36 72
 FT DOMAIN 79 158
 FT DOMAIN 159 172
 FT DOMAIN 173 434
 FT DISULFID 40 48
 FT DISULFID 42 60
 FT DISULFID 62 71
 FT DISULFID 162 296
 FT DISULFID 202 218
 FT DISULFID 210 285
 FT DISULFID 310 379
 FT DISULFID 342 358
 FT DISULFID 369 397
 FT ACT_SITE 217 217
 FT ACT_SITE 272 272
 FT ACT_SITE 373 373
 FT CARBOHYD 228 228
 SQ SEQUENCE 434 AA; 49400 MW; B0881048D66A55 CXC64;

Best Local Similarity 46.0%; Pred. No. 1.5e-72;
 Matches 196; Conservative 66; Mismatches 134; Indels 30; Gaps 8;
 QY 10 LCYLTVYS-DS-----KSGNELHQVPSNCDCLNGGVANCKKFSNTHMNCCKKKRGQGH 61
 DB 11 LCTLVGLDSDVYRQYKLSHKRPOHRECCCLNGGCTITFRFSQIKRLCLCEBGGLH 70
 QY 62 CEIDSKTCYEGNGHFGYKASDTWGRPCLPWSATVLO-QTYHAHRSALQGLGKH 120
 DB 71 CEIDTNSICVSGNGEDYRGAEDP-----GCLYNDHPSVIRMGDYHADLKNALQLGGLKH 126
 QY 121 YCNPNDRRRPWCYVGLKPLVQECMADCDGKSPSPPEELKQCCGKTARPRKII 180
 DB 127 YCNPNDRRRPWCYTK--RRYSIOE-----TPCSTIEECERTCGGRSFKYKIV 174
 QY 181 GGFETTENQPMFAIYRRHRSVYVYCGSGLISPCWVISATHCIFD----YPKEDYI 236
 DB 175 GGSQAEVETQPMAGIFQNIW-GTDQFLCGSGLIDPCWVLTAAHCFPIPKKQPNKSVYK 233
 QY 237 VYLGRSLNSNTQGEKFEVENLILKDYSDATLAHNDIALKIRSKRGCAQPSRTIQ 296
 DB 234 VFLGKSLTNDDEHQFVWDEILSHDPFDHTGNDNDIALIRIRASGCAVESNYVR 293
 QY 297 TILCPMTNDPQGTSCETIFGKENSIDYLYPEQLKMTYVYKLSHRECCQPRYSGEVT 356
 DB 294 TVCLPKRNLMLYDNTWCEIANGKONSIDYVYQRLMSATVNLISODCKKXYDSTRVT 353
 QY 357 TKMLCAADPQWKTDCGSGDGSLPGRMILTGVISWGRGALCKDPGYTVRVSHFL 416
 DB 354 DNWVCADPLMETDACKGDSGSPVVCENHGRMTLYGIVSGDGCANKNKGVITRYRYL 413
 QY 417 PWIRSH 422
 DB 414 NWIDSN 419
 RESULT 8
 ID1 URT2 DESRO STANDARD; PRT; 477 AA.
 AC P15638;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
 DE alpha-2) (BAT-PA) (T-plasminogen activator).
 OS Desmodus rotundus (Vampire bat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 CC Desmodontinae; Desmodus.
 OX NCBI_TaxId=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Salivary gland;
 RX MEDLINE=90036867; PubMed=2509450;
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
 RT "Isolation, characterization, and cDNA cloning of a vampire bat
 RT salivary plasminogen activator.";
 RL J. Biol. Chem. 264:17947-17952(1989).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,

Query Match

42.1%; Score 1009; DB 1; Length 434;

RA Goeddel D.V., Collen D.;
 RT "Cloning and expression of human tissue-type plasminogen activator
 RT cDNA in E. coli.";
 RT Nature 301:214-221(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=88262579; PubMed=3133640;
 RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
 RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
 RT from human fetal lung cells";
 RL Nucleic Acids Res. 16:5695-5695(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88054470; PubMed=2824147;
 RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
 RA Heising N.;
 RT "Expression of human uterine tissue-type plasminogen activator in
 RT mouse cells using BPV vectors.";
 RL DNA 6:461-472(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86196143; PubMed=3009482;
 RA Frieze Degen S.J., Rajput B., Reich B.;
 RT "The human tissue plasminogen activator gene,";
 RL J. Biol. Chem. 261:6972-6985(1986).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84298137; PubMed=6089198;
 RA NY T., Elgh F., Lund B.;
 RT "The structure of the human tissue-type plasminogen activator gene:
 RT correlation of intron and exon structures to functional and
 RT structural domains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86284200; PubMed=3090401;
 RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
 RA Opdenacker G., Volckaert G., Komabute W., Billiau A., Somer P.;
 RT "Cloning of cDNA coding for human tissue-type plasminogen activator
 RT and its expression in Escherichia coli.";
 RL Mol. Biol. Med. 3:279-292(1986).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOCORM SHORT).
 RC TISSUE=Umbilical vein;
 RX MEDLINE=90192129; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells.";
 RL Nucleic Acids Res. 18:1086-1086(1990).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strussberg R.L., Feingold R.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.T., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stanclevon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley R.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maitra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [9]
 RP SEQUENCE OF 212-361 FROM N.A.
 RX MEDLINE=83169656; PubMed=6572897;
 RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
 RA Josephson S.;
 RT "Isolation of cDNA sequences coding for a part of human tissue
 RT plasminogen activator.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
 RN [10]
 RP SEQUENCE OF 1-36 FROM N.A.
 RX MEDLINE=85289338; PubMed=3161893;
 RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
 RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region.";
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [11]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1368681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [12]
 RP SEQUENCE OF 36-562.
 RC TISSUE=Melanoma;
 RX MEDLINE=85000468; PubMed=6433976;
 RA Pohl G., Kaelin-Lang A., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences.";
 RL Biochemistry 23:3701-3707(1984).
 RN [13]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE=Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator.";
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [14]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells.";
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [15]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain.";
 RL Biochemistry 30:2311-2314(1991).
 RN [16]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaaskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in Escherichia coli.";
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Landa D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator.";
 RL J. Mol. Biol. 258:117-135(1996).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;

DR SMART; SMO0020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS0186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1
 FT CHAIN 37
 FT NCB1 431
 FT DOMAIN 37
 FT DOMAIN 82
 FT DOMAIN 179
 FT ACT_SITE 226
 FT ACT_SITE 275
 FT ACT_SITE 382
 FT DISULFID 41
 FT DISULFID 46
 FT DISULFID 63
 FT DISULFID 65
 FT DISULFID 74
 FT DISULFID 82
 FT DISULFID 103
 FT DISULFID 134
 FT DISULFID 158
 FT DISULFID 168
 FT DISULFID 211
 FT DISULFID 219
 FT DISULFID 288
 FT DISULFID 313
 FT DISULFID 345
 FT DISULFID 378
 FT CARBOHYD 379
 FT CARBOHYD 352
 FT CARBOHYD 352
 SQ SEQUENCE 431 AA; 48221 MW; 699BSB675B162CBF CMC64;
 Query Match 36.2%; Score 867.5; DB 1; Length 431;
 Best Local Similarity 41.6%; Pred. No. 2.4e-61;
 Matches 185; Conservative 61; Mismatches 152; Indels 47; Gaps 10;

TPA_RAT ID TPA_RAT STANDARD; PRT; 559 AA.
 AC P16637;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (cPA)
 DE (c-Pa) (c-plasminogen activator).
 GN PLAT.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RX MEDLINE=9170114; PubMed=3148445;
 RA Ny T., Leonardson G., Heuvel A.J.W.,
 RT "Cloning and characterization of a cDNA for rat tissue-type
 RT plasminogen activator.";
 RL DNA 7:671-677(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90130448; PubMed=2105315;
 RA Feng P., Ohlsson M., Ny T.,
 RT "The structure of the TATA-less rat tissue-type plasminogen activator
 RT gene. Species-specific sequence divergences in the promoter predict
 RT differences in regulation of gene expression.";
 RL J. Biol. Chem. 265:2022-2027(1990).
 CC -1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
 CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
 CC controlling plasmin-mediated proteolysis, it plays an important
 CC role in tissue remodeling and degradation, in cell migration and
 CC many other physiological events.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
 CC bond.
 CC -1- PTM: SUBCELLULAR LOCATION: Secreted; extracellular.
 CC -1- PUBMED: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
 CC chain. Binding to fibrin enhances its catalytic activity.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 2 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M23697; AAA41812.1; -;
 DR EMBL; M31197; AAA42261.1; -;
 DR EMBL; M31185; AAA42261.1; JOINED.
 DR EMBL; M31186; AAA42261.1; JOINED.
 DR EMBL; M31187; AAA42261.1; JOINED.
 DR EMBL; M31188; AAA42261.1; JOINED.
 DR EMBL; M31189; AAA42261.1; JOINED.
 DR EMBL; M31190; AAA42261.1; JOINED.
 DR EMBL; M31191; AAA42261.1; JOINED.
 DR EMBL; M31192; AAA42261.1; JOINED.
 DR EMBL; M31193; AAA42261.1; JOINED.
 DR EMBL; M31194; AAA42261.1; JOINED.
 DR EMBL; M31195; AAA42261.1; JOINED.
 DR EMBL; M31196; AAA42261.1; JOINED.
 DR EMBL; A19618; CAA01482.1; -;
 DR PIR; A35029; A35029.
 DR HSSP; P00750; IRTF.

```
Query March      35.4%; Score 847.5; DB 1; Length 559;  
Beet Local Similarity   36.7%; Pred.No.12e-59;  
Matches 183; Conservative    65; Indels    97; Gaps    12;  
  
QY      18 SKGSNELHQVF-----SNDCCLNGRCVSNKPKFSNTHMCNCPKKPFGGQHCEIDSKTCYEG 73  
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
```

DB	URL1	DESCRO	STANDARD	PRT	477	AA
DB	69	NSGLAQCSVPRVSCSEBPCFCNGCTCOALVYPSDF-VCCCPDGFVGRKCIDTRATCFEG				127
QY	74	NGHFRGCAISIDTMRPCLPANAVALVLOQTVAHNSDALQLGLGKNNYCRNPNNRRPWC				133
DB	128	QGITTRGWTSAENAGCACINNNSSALSKPYSARRPNAIKGLGNNYCRNPDRDYKPMC				187
QY	134	VYQVGLKFLVQECM-----VHDCADGK-----KPSPP-----161				
DB	188	VYFRAKGTTFEFCSTPACPKGPTEDCYGKVTVYRGTHSFTTSSKASCLPNNSMILGKTY				247
QY	162	-----BELKFO-----CGOKTLR-PRF				177
DB	248	TAMRANSQALGLGRNHYCRNPDPGAKPMCHVWKDKLTWEYCDMSPCSTCGLRQYKQPOF				307
QY	178	KIIGGEFTTINQPPFAIY-RRHGGSVTVYCGSGLISPCVAVSATHCFIDYPKKEDI				236
DB	308	RIKGGLFTDITSHPMQAAIFVNNKRSRGRFRFCGVLLISCVLSAAMCFVEFRPHHLK				367
QY	237	VYLGSRANSTQCGMKFEVENTLHKQYSAPTLHAHNDIALKTRSKRCGRCAQSPRTIQ				296
DB	368	VVLGTVVVPGEESQTFEIEIKYIHKAFDDT--YNDIALQLRSDSSQCAQESSVYG				425
QY	297	TICLPSMYNDPOF-----GTSCEITFGKENSTDYLPBQLKMTVVLISHRECQPHYYG				352
DB	426	TACLP-----DDVQVLPDMTECELSGTGKHEASSPPFSDLKANHRLVYSSRTSGHLFN				481
QY	353	SEVTTKMLCAADP-----QWKTDSCQSGSGPLVCSLOGRMVLTGVISNGRGCAIXDKPG				407
DB	482	KTITSNMLCAGDTRTGQNDVHDA CQDSGGLVCMIDRMVTLGIISWGLGCGQKDVPG				541
QY	408	VYTRVSHFLPWIRSHITKE				425
DB	542	IYTKVTNYLNMVIOQNMKQ				559

RESULT 12

URL1	DESCRO	STANDARD	PRT	477	AA
AC	P98119;				
DT	01-FEB-1996 (Rel. 33, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA alpha-1).				
OS	Desmodus rotundus (Vampire bat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;				
OC	Desmodontinae; Desmodus.				
OX	NCBI_TaxID=9430;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Salivary gland;				
FX	MEDLINE=92039036; PubMed=1937019;				
RA	Kreutzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,				
RA	Alagon A., Donner P., Schleuning W.D.;				
RT	"The plasminogen activator family from the salivary gland of the				
RT	vampire bat Desmodus rotundus: cloning and expression.";				
RL	Gene 105:229-237(1991).				
RN	[2]				
RP	CHARACTERIZATION.				
FX	MEDLINE=93393059; PubMed=1309059;				
RA	Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,				
RA	Kreutzschmar J., Haendler B., Langer G., Baidou B., Wilt W.,				
RA	Donner P.;				
RT	"Plasminogen activators from the saliva of Desmodus rotundus (common				
RT	vampire bat): unique fibrin specificity.";				
RL	Ann. N.Y. Acad. Sci. 667:395-403(1992).				
RN	[3]				
RP	X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).				
RC	TISSUE=Salivary gland;				
FX	MEDLINE=98022741; PubMed=9354616;				
RA	Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,				
RA	Schleuning W.D., Bode W.;				


```

Db 193 VIRAGFTSSCSVPVCS-----KATGCLNKKEPOLHSTGCLFTDITSHPWQ 240
QY 194 AAIYRRHRRGGS-VYVCGGSLISPCWVATSHCFIDYPKKEDI-----VYLGSRNSN 247
Db 241 AAIIPAQRBRSSGGERFLCGGILISGCVLTANHC-----GESYLPDLKVLGRTYVKP 295
QY 248 TQGBMKFEVENLIHKDYSADTLAHNDIALKLRISREGCAQPSRTIQTCLPSMYNDP 307
Db 296 GEEBQFVKVKKYIYHKEFDDT--YNNDIALLQLKSDSPQCAQPSDSVRAICLPEANLQL 353
QY 308 QFETSCITTFPGKENSNDIYLPBOLKTVYKLSHRECCQPHYGVTSVTKMLCAADPQW 367
Db 354 PDWTECLLSYGRKSSSPFYSBQLKGVHLYPSSRCAPKFLPKTVTNMCAQDTRS 413
QY 368 KT-----DSCQGDSPGLVCSLQGRWTLGIVSWGRGCLKDKPGYTRVSHPLPIRS 421
Db 414 GEIYPNVHDCQGDSPGLVCMNDNMTLGIISMGVCGEKDVPYVTVNTLGHIRD 473
QY 422 H 422
Db 474 N 474

RESULT 13
TPA_MOUSE STANDARD; PRT; 559 AA.
AC P1214; 091VP2;
DT 01-JUL-1989 (Rel. 11, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tpa)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86087303; PubMed=2826484;
RA Rickles R.J., Barrow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during F9 teratocarcinoma cell
RT differentiation."
RL J. Biol. Chem. 263:1563-1569(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Fellinold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Shuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Blatchenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toohyuki S., Cavani P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
CC controlling plasmin-mediated proteolysis, it plays an important
CC role in tissue remodeling and degradation, in cell migration and

```

```

CC many other physiopathological events.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
CC bond.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-108 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
CC chain. Binding to fibrin enhances its catalytic activity.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, J03520; AAA0470.1; -.
DR EMBL, BC011256; AAH11256.1; -.
DR PIR, A29941; A29941.
DR HSP, P00750; 1ASH.
DR MEROPS, S01.232; -.
DR K0D, MGI:97610; Plat.
DR InterPro, IPR009003; Cys_Ser_trypsin.
DR InterPro, IPR006209; EGF-like.
DR InterPro, IPR000083; Fibronctn1.
DR InterPro, IPR006210; IEGF.
DR InterPro, IPR000001; Kringle.
DR InterPro, IPR001254; Peptidase_S1.
DR InterPro, IPR001314; Peptidase_S1A.
DR Pfam, PF00008; EGF_1.
DR Pfam, PF00039; fn1_1.
DR Pfam, PF00051; kringle_2.
DR Pfam, PF00089; trypsin_1.
DR PRINTS, PR00722; CHYMOTRYPSIN.
DR PRINTS, PR00018; KRINGLE.
DR ProDom, PD000395; Kringle_2.
DR SMART, SM00181; EGF_1.
DR SMART, SM00058; FN1_1.
DR SMART, SM00130; KR_2.
DR SMART, SM00020; TRYP_SPC_1.
DR PROSITE, PS00022; EGF_1; 1.
DR PROSITE, PS01186; EGF_2; 1.
DR PROSITE, PS50026; EGF_3; 1.
DR PROSITE, PS01253; FIBRONECTIN_1; 1.
DR PROSITE, PS00021; KRINGLE_1; 2.
DR PROSITE, PS50070; KRINGLE_2; 2.
DR PROSITE, PS50240; TRYPsin_DOM; 1.
DR PROSITE, PS00134; TRYPsin_HIS; 1.
DR PROSITE, PS00135; TRYPsin_SER; 1.
KW Plasminogen activator; Hydrolyase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT_SITE 355 355
FT ACT_SITE 404 404

```

```

FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 G-> A (IN REF. 1).
FT CONFLICT 260 260 P-> A (IN REF. 1).
FT CONFLICT 325 325
SQ SEQUENCE 559 AA; 63122 MM; 8CCE2BD94514D9 CRC64;

Query Match 34.6%; Score 828.5; DB 1; Length 559;
Best Local Similarity 35.5%; Pred. No. 3.9e-58;
Matches 177; Conservative 66; Mismatches 158; Indels 97; Gaps 10;

18 SKGSNELHOVP-----SNCDINGSTCVSNKYFSNIHMCNPKKFGQCHETDKSKTCYEG 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db NSGLVQCHSVPRSCSEPRCFNGGTCQALYFSDP--VCQCPDGFVGRKCDIDTRATCFEE 127

74 NGHYTRKGAASDITMGPRCLPMNSATVLOQTHARSDLOGLGKHYCRAPDNRARRPWC 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db QGITRRGWSYAESGAEICINNSSVLSLKPYNARPNAIKLGANHCNRPDRLKPMC 187

128 YVQGLKLPVOECMWHDCADGKKPS----- 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY YVQGLKLPVOECMWHDCADGKKPS----- 158

134 YVQGLKLPVOECMWHDCADGKKPS----- 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY YVQGLKLPVOECMWHDCADGKKPS----- 158

188 YVFRAGKYTEFCSTPACPKKSEBCYGVKGYRGTHTSLTTSQASCLPMNSIVLMGKSY 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db YVFRAGKYTEFCSTPACPKKSEBCYGVKGYRGTHTSLTTSQASCLPMNSIVLMGKSY 247

159 -----SPEELK-----FOCG-QKTLRPF 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY -----SPEELK-----FOCG-QKTLRPF 177

248 TAWRTNSQALGLGRHNYCRNPDGDARPMCHWKDKLTWEYCDMSPCSTCGLRQYKRPQF 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db TAWRTNSQALGLGRHNYCRNPDGDARPMCHWKDKLTWEYCDMSPCSTCGLRQYKRPQF 307

178 KIIGEFETIENQPFPAIY--RRHSGSVTVVCGSLISPCWVTSATTCFIDYRKEDI 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY KIIGEFETIENQPFPAIY--RRHSGSVTVVCGSLISPCWVTSATTCFIDYRKEDI 236

308 RIKGGLYDITSHPWQAPLIFYKNKRSPEGRFLCGSVLISCVLSAAHCFIERPPNHLK 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db RIKGGLYDITSHPWQAPLIFYKNKRSPEGRFLCGSVLISCVLSAAHCFIERPPNHLK 367

237 VYLGSRINSTQGEKKEVENLILHKDYSADTLAHNDIALILKIRSEKGRCAOPSRITQ 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY VYLGSRINSTQGEKKEVENLILHKDYSADTLAHNDIALILKIRSEKGRCAOPSRITQ 296

368 VVLGRTYAVPGESEOTFEIKYIYHEFFDDT--YDNDIALLOQRSQSKCAQSSSVG 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db VVLGRTYAVPGESEOTFEIKYIYHEFFDDT--YDNDIALLOQRSQSKCAQSSSVG 425

237 TICLPSTMYNDPQF-----GTSCEITSGKENSTDIYYPOLKMTYVKLISHRECQPHYG 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY TICLPSTMYNDPQF-----GTSCEITSGKENSTDIYYPOLKMTYVKLISHRECQPHYG 352

426 TACLP-----DPMQLPDWTECELSGCGKHEASPPFSRLKKAHYRLKLPSSRCTSOHFN 481
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db TACLP-----DPMQLPDWTECELSGCGKHEASPPFSRLKKAHYRLKLPSSRCTSOHFN 481

353 SEVTKMLCAADP-----QMKTDSQCGSGGFLVCSLQGRMTLTGIVSMGSGCALKDXG 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY SEVTKMLCAADP-----QMKTDSQCGSGGFLVCSLQGRMTLTGIVSMGSGCALKDXG 407

482 KTVINNMCAQGTSGGNQDLHDAQQDSGGFLVCMINKOMTLTGIIISWGLGCGQKDPFG 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db KTVINNMCAQGTSGGNQDLHDAQQDSGGFLVCMINKOMTLTGIIISWGLGCGQKDPFG 541

408 VYTRVSHFLPWIRSHSTKE 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY VYTRVSHFLPWIRSHSTKE 425

542 VYTKVTNYLDWTHDMKQ 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db VYTKVTNYLDWTHDMKQ 559

RESULT 14
TPA_BOVIN STANDARD; PRT; 566 AA.
AC 028198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)

```

```

DE (t-PA) (t-Plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -I- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
controlling plasmin-mediated proteolysis, it plays an important
role in tissue remodeling and degradation, in cell migration and
many other physiological events.
CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -I- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
bond.
CC -I- SUBCELLULAR LOCATION: Secreted; extracellular.
CC -I- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN, TISSUE KALIKEIN OR FACTOR XA.
CC -I- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
chain. Binding to fibrin enhances its catalytic activity.
CC -I- SIMILARITY: Belongs to peptidase family S1.
CC -I- SIMILARITY: Contains 1 EGF-like domain.
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
CC -I- SIMILARITY: Contains 2 kringle domains.
CC -I- SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sb-sib.ch).
CC -----
DR EMBL; X85800; CA59795.1; -.
DR HSBP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006083; Fibnctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF. 1.
DR Pfam; PF00039; fn1. 1.
DR Pfam; PF00051; kringle. 2.
DR Pfam; PF00089; trypsin. 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR0018; KRINGLE.
DR ProDom; PD000395; Kringle. 2.
DR SMART; SMO0181; EGF. 1.
DR SMART; SMO0058; FN1. 1.
DR SMART; SMO0130; KR. 2.
DR SMART; SMO0020; TRYD_SPE. 1.
DR PROSITE; PS00022; EGF_1. 1.
DR PROSITE; PS01186; EGF_2. 1.
DR PROSITE; PS50026; EGF_3. 1.
DR PROSITE; PS01253; FIBRONECTIN_1. 1.
DR PROSITE; PS00021; KRINGLE_1. 1.
DR PROSITE; PS50070; KRINGLE_2. 2.
DR PROSITE; PS50240; TRYPSIN_DOM. 1.
DR PROSITE; PS00134; TRYPSIN_HIS. 1.
DR PROSITE; PS00135; TRYPSIN_SER. 1.
DR Plasmidogen activation; Hydrolyase; Serine protease; Glycoprotein;
KM

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 117.888 Seconds
(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-2

Perfect score: 2394
Sequence: 1 MALLARLLLCVIVSDSKG.....VSHFLPWIRSHTEKENGALL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvrius:*
16: sp_bacteriophage:*
17: sp_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1992	83.2	433	6	Q8MIL0	Q8MIL0 oryctolagus
2	1987	83.0	433	6	Q8MIL7	Q8MIL7 oryctolagus
3	982.5	41.0	214	6	Q9XU70	Q9XU70 oryctolagus
4	882	36.8	231	11	Q8C6L2	Q8C6L2 mus musculus
5	871	36.4	154	4	Q968E8	Q968E8 homo sapien
6	870.5	36.4	154	4	Q9BU99	Q9BU99 homo sapien
7	868.5	36.3	562	4	Q86YK8	Q86YK8 homo sapien
8	826	34.5	564	6	Q8MK31	Q8MK31 oryctolagus
9	822.5	34.4	562	6	Q8SG23	Q8SG23 sus scrofa
10	784.5	32.8	395	4	Q9B2M1	Q9B2M1 homo sapien
11	731.5	30.6	540	13	Q800Y7	Q800Y7 meleagris g
12	727	30.4	653	11	Q8VCS4	Q8VCS4 mus musculus
13	693.5	29.0	616	6	Q97507	Q97507 sus scrofa
14	689	28.8	615	4	Q8IZ25	Q8IZ25 homo sapien
15	680.5	28.4	597	11	Q35727	Q35727 mus musculus
16	680.5	28.4	609	11	Q80YCS	Q80YCS mus musculus

17	677	28.3	157	6	Q9TV48	Q9TV48 bos taurus
18	674.5	28.2	517	11	Q8K0D2	Q8K0D2 mus musculus
19	661.5	27.6	560	4	Q14520	Q14520 homo sapien
20	585	24.4	128	6	Q97587	Q97587 oryctolagus
21	498	20.8	810	4	Q15146	Q15146 homo sapien
22	496	20.7	454	6	Q46506	Q46506 papio hamad
23	496	20.7	812	11	Q9R0W3	Q9R0W3 rattus norv
24	495.5	20.7	300	4	Q96EP3	Q96EP3 homo sapien
25	495	20.7	103	6	Q9SM89	Q9SM89 equus caball
26	491.5	20.5	429	13	Q8AVB0	Q8AVB0 brachydanto
27	482.5	20.2	334	6	Q46507	Q46507 papio hamad
28	472.5	19.7	806	6	Q18783	Q18783 macropus eu
29	463.5	19.4	868	5	Q9YIV3	Q9YIV3 polyandroca
30	442.5	18.5	327	4	Q8N171	Q8N171 homo sapien
31	435.5	18.2	284	4	Q8NFB6	Q8NFB6 homo sapien
32	433.5	18.1	761	11	Q99J08	Q99J08 rattus norv
33	426.5	17.8	267	5	Q9BK47	Q9BK47 ludia foli
34	426	17.8	558	4	Q86YK4	Q86YK4 homo sapien
35	425.5	17.8	505	5	Q966V4	Q966V4 halocynthia
36	424	17.7	537	4	Q9BYE1	Q9BYE1 homo sapien
37	420	17.5	471	11	Q8CFE0	Q8CFE0 mus musculus
38	420	17.5	581	4	Q9BYE2	Q9BYE2 homo sapien
39	418.5	17.5	277	11	Q80WM7	Q80WM7 mus musculus
40	417.5	17.4	425	13	Q804X7	Q804X7 gallus gall
41	414	17.3	638	11	Q8R0P5	Q8R0P5 mus musculus
42	411.5	17.2	276	11	Q8CGR6	Q8CGR6 mus musculus
43	409	17.1	371	11	Q8CJ16	Q8CJ16 rattus norv
44	409	17.1	445	11	Q8CJ17	Q8CJ17 rattus norv
45	405.5	16.9	422	4	Q8WYCI	Q8WYCI homo sapien

ALIGNMENTS

RESULT 1	Q8MIL0	PRELIMINARY;	PRT;	433 AA.
ID	Q8MIL0			
AC	Q8MIL0			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Urokinase-type plasminogen activator.			
GN	PLAU.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_Taxid=9986;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22155945; PubMed=12149463;			
RA	Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,			
RA	Dichek D.A.,			
RT	"Increased expression of urokinase during atherosclerotic lesion			
RT	development causes arterial constriction and lumen loss, and			
RT	accelerates lesion growth."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
DR	EMBL; AY122285; AM83187.1; -			
DR	GO; GO:004263; F:chymotrypsin activity; IEA.			
DR	GO; GO:0016301; F:kinase activity; IEA.			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	GO; GO:0004295; F:trypsin activity; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR009003; Cys_Ser_trypsin.			
DR	InterPro; IPR006209; BGF_1like.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR001254; Peptidase S1.			
DR	InterPro; IPR001314; Peptidase S1A.			
DR	InterPro; IPR008293; Pept_S1A_TPA.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			

PRINTS; PRO0018; KRINGLE.
 DR Prodom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS00022; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR PIRSF; PIRSF001144; Utk_Plaam_act; 1.
 DR Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
 KW SEQUENCE 433 AA; 4875 MW; 65E64F3641554990 CRC64;
 SO

Query Match 83.2%; Score 1992; DB 6; Length 433;
 Best Local Similarity 83.4%; Pred. No. 1.6e-182;
 Matches 361; Conservative 26; Mismatches 44; Indels 2; Gaps 1;

1 MRALLARLLLCVAVVSDSKSGNELHOV--PSNCDCLANGCTVSNKYFSNIHMCNCPKKG 58
 1 MRVLLVCLLCALVVDSESGSHELHGVSDASNCGCLNGCTVYKXFSNIHMCNCPKKG 60
 59 GQCEIDKSKTCYENGHFFRGKASTDTMGRPCLPNSATVLCQTYHAHSDALQGLGK 118
 61 GEHCEIDTLKTCYHGDHSTRGKANTDIMORPCLAMNSANVLTXYHAHSDALQGLGK 120
 119 HNYCRNPDNRPRMPCYVGLKPLVOCMWHDCADKKSPPEELKFGCGQKTLRPRFK 178
 121 HNYCRNPDHRRPRMPCYVGLKPLVOCMWHDCADKKSPPEELKFGCGQKTLRPRFK 180
 179 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 238
 181 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 240
 239 LGRSLNSNTOGEMKFEVENLIHKDYSADTLAHNDIALIKIRSKGRCAQPSRTIQT 298
 241 LGRSLNSMTGEMKFEVEOLILHGYADTLAHNDIALIKIRSKGRCAQPSRTIQT 300
 299 CLPSTMYNDPQFGTSCEITGFGKENSVDYLPBOLKMTVVKLISHRCQCPHYGSEVTTK 358
 301 CLPSTMYNDPQFGTSCEITGFGKENSVDYLPBOLKMTVVKLISHRCQCPHYGSEVTTK 360
 359 MLCADPQWKTDSCQGSGLPVCISLOGRMTLTGIVSWGRCALKDKRPVYTRVSHFLPW 418
 361 MLCADPQWKTDSCQGSGLPVCISLOGRMTLTGIVSWGRCALKDKRPVYTRVSHFLPW 420
 419 IRSHTKEENGAL 431
 421 IRSHTKEENGAL 433

RESULT 2
 Q8WHY7 PRELIMINARY; PRT; 433 AA.
 AC Q8WHY7;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Urokinase-Type plasminogen activator.
 GN UROKINASE.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCB1_TaxId=9986;
 OX NCB1_TaxId=9986;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yano W., Watanabe M.;
 RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
 RT complete cde.";

Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AY029517; AK40239.1; -
 DR EMBL; AB087224; BAC02685.1; -
 DR GO; GO:004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:008233; F:peptidase activity; IEA.
 DR GO; GO:004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR008293; Pept_S1A_UPA.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR PRINTS; PRO0018; KRINGLE.
 DR Prodom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR PIRSF; PIRSF001144; Utk_Plaam_act; 1.
 DR Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
 KW SEQUENCE 433 AA; 4844 MW; 6D35A571010A6E CRC64;
 SO

Query Match 83.0%; Score 1987; DB 6; Length 433;

Best Local Similarity 83.1%; Pred. No. 4.8e-182;
 Matches 360; Conservative 26; Mismatches 45; Indels 2; Gaps 1;

1 MRALLARLLLCVAVVSDSKSGNELHOV--PSNCDCLANGCTVSNKYFSNIHMCNCPKKG 58
 1 MRVLLVCLLCALVVDSESGSHELHGVSDASNCGCLNGCTVYKXFSNIHMCNCPKKG 60
 59 GQCEIDKSKTCYENGHFFRGKASTDTMGRPCLPNSATVLCQTYHAHSDALQGLGK 118
 61 GEHCEIDTLKTCYHGDHSTRGKANTDIMORPCLAMNSANVLTXYHAHSDALQGLGK 120
 119 HNYCRNPDNRPRMPCYVGLKPLVOCMWHDCADKKSPPEELKFGCGQKTLRPRFK 178
 121 HNYCRNPDHRRPRMPCYVGLKPLVOCMWHDCADKKSPPEELKFGCGQKTLRPRFK 180
 179 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 238
 181 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 240
 239 LGRSLNSNTOGEMKFEVENLIHKDYSADTLAHNDIALIKIRSKGRCAQPSRTIQT 298
 241 LGRSLNSMTGEMKFEVEOLILHGYADTLAHNDIALIKIRSKGRCAQPSRTIQT 300
 299 CLPSTMYNDPQFGTSCEITGFGKENSVDYLPBOLKMTVVKLISHRCQCPHYGSEVTTK 358
 301 CLPSTMYNDPQFGTSCEITGFGKENSVDYLPBOLKMTVVKLISHRCQCPHYGSEVTTK 360
 359 MLCADPQWKTDSCQGSGLPVCISLOGRMTLTGIVSWGRCALKDKRPVYTRVSHFLPW 418
 361 MLCADPQWKTDSCQGSGLPVCISLOGRMTLTGIVSWGRCALKDKRPVYTRVSHFLPW 420
 419 IRSHTKEENGAL 431
 421 IRSHTKEENGAL 433

RESULT 3
 Q9XT70 PRELIMINARY; PRT; 214 AA.
 ID Q9XT70

```

AC 09XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin Y., Idell S.;
RT "Partial mRNA of rabbit uPA."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; A039351.1; -.
DR HSSP; P00749; 1EUN.
DR MEROPS; S01.231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR GlycoProtein; Hydroxylase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 214 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C3B0D7E CRC64;

Query Match 41.0%; Score 982.5; DB 6; Length 214;
Best Local Similarity 82.8%; Pred. No. 5.3e-86;
Matches 178; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

QY 87 MGRPCLPMSATVLTQTYAHRSDALQGLGKINYCNPDRRRPWCYVGVGKPLVQEC 146
DB 1 MDRPCLAMNSANVLTQTYAHRSDALQGLGKINYCNPDRRRPWCYVGVGKPLVQEC 60

QY 147 MNRDCAADGKRPSPPELAKQCCQKTLRRFKIIGSEFTTIENQWFAIYRRHSGSVT 206
DB 61 KVRD-SSGKKRPALPGLLEFQCCQKALRPFRKIIGSEFTTIENQWFAIYRRHSGSVT 119

QY 207 YVCGSLISPCWYISATHCFIDYPKKEDYIVYGRSLNNTGQEMKFEVENTILHKDYS 266
DB 120 YVCGSLISPCWYISATHCFIDYPKKEDYIVYGRSLNNTGQEMKFEVENTILHKDYS 179

QY 267 ADTLAHHNDIALKIRSKRGCAQPSRTIQTICLP 301
DB 180 ADTLAHHNDIALKIRSKRGCAQPSRTIQTICLP 214

RESULT 4
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC 08C6L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK054349; BAC35743.1; -.
DR PIR; P70534; P70534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
FT NON_TER 231 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 36.8%; Score 882; DB 11; Length 231;
Best Local Similarity 67.7%; Pred. No. 2.6e-76;
Matches 155; Conservative 27; Mismatches 45; Indels 2; Gaps 2;

QY 1 MRLLARLLCVLVSDSKGSENLHGV-PSNCDLNGTCVSNKYSTNHWNCQPKKFGG 59
DB 1 MKWVLASLFICALLVNKGSEGVLAGAPDSNCCQNGCVSVKYSRIRRCSCPRKFOG 60

QY 60 QHCEIDKSTCYGNGNHFYRGRKASTDTMTGRPCLPMSATVLTQTYAHRSDALQGLGKH 119
DB 61 EHEIDASKTCYHGNDSYRGRKANTDTKGRPCLAMNAPALQKPYAHRSDALSLGKGK 120

QY 120 NYCRNPDRRRPWCYVGVGKPLVQECMVDCAHGKSPPELAKFOGQKTLRRFKI 179
DB 121 NYCRNPDRRRPWCYVGVGKPLVQECMVDCAHGKSPPELAKFOGQKTLRRFKI 180

QY 180 IGSEFTTIENQWFAIYRRHSGS-VTVVCGSLISPCWYISATHCFI 227
DB 181 VGSEFTEVENQWFAIYQKNGKSGSPSPFCGSLISPCWYISATHCFI 229

RESULT 5
Q96SE8 PRELIMINARY; PRT; 154 AA.
AC Q96SE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human
RT urokinase-type plasminogen activator."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```


Bai X., Fu J., Wang W., Xi X., Ruan C.;
 RT "overexpression of the amino-terminal fragment of human urokinase-type
 RT plasminogen activator in breast cancer cells results in decreased
 RT tumor invasion, growth and angiogenesis."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL: AY029537; AAK8734.1; -.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR Pfam: PF00051; kringle; 1.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 KM Glycoprotein; Kinase; Kringle.
 SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF505572 CRC64;

Query Match 36.4%; Score 871; DB 4; Length 154;
 Best Local Similarity 99.4%; Pred. No. 1.8e-75;
 Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRALLALLLCVLYVSDSKSGSNEHGVPSNDCCLNGSTCVSNKYFSNIHMCNCKKRGQ 60
 DB 1 MRALLALLLCVLYVSDSKSGSNEHGVPSNDCCLNGSTCVSNKYFSNIHMCNCKKRGQ 60
 QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGKPCLPNNSATVLOQTTHARSDALQGLGKN 120
 DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGKPCLPNNSATVLOQTTHARSDALQGLGKN 120
 QY 121 YCRNPDRRRPWCYVQGLKPLVQECWVHDCADG 154
 DB 121 YCRNPDRRRPWCYVQGLKPLVQECWVHDCADG 154

RESULT 6

Q9BU99 PRELIMINARY; PRT; 516 AA.
 AC Q9BU99;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL: BC002795; AA02795.1; -.
 DR HSSP; P00750; IASH.
 DR GO: GO:0004263; F:chymotrypsin activity; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.

DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 2.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 2.
 DR SMART: SM00020; TRYP_Spc; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 2.
 DR PROSITE: PS00070; KRINGLE_2; 2.
 DR PROSITE: PS00024; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 KM EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KM Serine protease.
 SQ SEQUENCE 516 AA; 57370 MW; BAB1901FDC96800 CRC64;

Query Match 36.4%; Score 870.5; DB 4; Length 516;
 Best Local Similarity 37.8%; Pred. No. 9.2e-75;
 Matches 195; Conservative 57; Mismatches 165; Indels 99; Gaps 12;

QY 2 RALLAALLLC-VLYVSDSKSGSNEH-----QVPSNDCCLNGSTCVSNKYFSNIH 50
 DB 6 RGLCVLLLCGAVFVPSQ---EIHARFRGARASYGCSBPRCFNGCTOQALYFSDP-V 61
 QY 51 CNCPKFFGQHCIEDSKTCYEGNGHFYRGKASTDTMGKPCLPNNSATVLOQTTHARSD 110
 DB 62 CQCPBGFARKCCGIDIRATCYEDQGISYRGWSTASGAECTWMSALAKPYSGRRPD 121
 QY 111 ALQGLGKNHYCNPNRRPWCYVQGLKPLVQECWVHDCADG----- 154
 DB 122 AIRLGANNHYCNPNRRPWCYVQGLKPLVQECWVHDCADG----- 181
 QY 155 -----KKPS-----SPPRELK----- 165
 DB 182 SLTESGASCLPNSMILIGVYTAQNPISAQALGKHNCRANDGAKPRCHVLKRRRL 241
 QY 166 -----FOCG-OKTLRPRFKIIGEFITINOPWFAIYRRH-RGGSVTVYCGSLI 214
 DB 242 WEYCDVPSCSTGCLRGYSQGFRIKGLFADIAHPQAAIIFAKHRRSPDERLCCGILL 301
 QY 215 SPCWVISANTCFIDYPKEDYIYLGSRSLNSTOEGMKREVENLLIHQYSADTLAHNN 274
 DB 302 SSCMIISAAHCPOERPPPHLTVILGRYRVVGESEQKFEVKKIVHKEFDDT--YDN 359
 QY 275 DIALTKRSKEGCAOPSRTIOTICLPWYNDRPOFGSCETIFGKENSTDYLYPEQLKM 334
 DB 360 DIALQKSDSSKCAQESSVVRIVCLPPADLQLPDWTCELSGKGHSLSPYSRLKE 419
 QY 335 TVVKLISHRECQOPHYGSEVTTKMLCAAD-----PQWKT-DSQQDGGPVLCSLGRM 388
 DB 420 AHVRLVSSRCTSQHLNLRVTVDNMLCAGDTRSGSPQANLHDACQSGSGPLVCLNDGRM 479
 QY 389 TLVGIYSWGRGCAIKDKPGYTRYVSHLPWIRSHTK 424
 DB 480 TLVGIISWGLGCGQKDVPGYTRYVTVLWIRUNMR 515

RESULT 7

Q86YK8 PRELIMINARY; PRT; 562 AA.
 AC Q86YK8;
 DT 01-JUN-2003 (TREMblrel. 24, Created)
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Tissue plasminogen activator.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP Liu Y., Xu L., Zeng Y., He X.;

```

RT "CDNA of tissue plasminogen activator."
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY221101; AAC34406.1;
DR GO; GO:0005576; F:chymotrypsin activity; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 562 AA; 62902 MW; 837D98392F6EDDIF CRC64;

Query Match 36.3%; Score 868.5; DB 4; Length 562;
Best Local Similarity 37.9%; Pred. No. 1.6e-74;
Matches 187; Conservative 57; Mismatches 160; Indels 89; Gaps 11;

QY 18 SKGSNELHQVP--SNCD---CLANGTCVSNKYFSNIMHCNPKKFGQCHCEIDSKTCYEG 73
DB 72 NSGRAGQHSVPVASCSEPRFCNGTCCQALYFSDP--VCQCPREGAGKCEIDTRATCYED 130
QY 74 NGHFYRGKASTDTWGRPCLPMSATVLOQTYHAHRSDALQGLKKNYCNPNDRRRPWC 133
DB 131 QGISYRGWTSTAESGAECTWMSALAQKPYSGRRPDALRLGLGNHNYCNPNDRDSKPC 190
QY 134 YVVGGLKELVQECVNHQCADG----- 154
DB 191 YVERAGKYSSEFCSTPACSSGNSDCYFGNSAVRGTHSLTESGASCLPMSMILIGNVT 250
QY 155 -KKPS-----SPEELK-----PQCG--OKTLRPRFK 178
DB 251 AQNPBSAALQGLKKNYGRNDGAKPWCCHYLKXRLTWECYDVVSCSTCGLRQVSGQFR 310
QY 179 IIGGEFTIENQWPAFYLRH--RGGSVTVVCGSLSPCWVSIATHCFTIDYPRKEDYIV 237
DB 311 IKGGLFADISHMPQALIFAKHRRSPERFLCGIILISSCWIIISAACHFQERFPHLTV 370
QY 238 YLGRSRINSTQGMKEVENLILKDYASDTLAHNDIALKTRSGKCAQPSRTIOT 297
DB 371 ILKRTYRVVGEEOKEVEKYVHKEFDDT--YNDIDIALLOKSSSCADQSSVVRT 428
QY 298 ICLPSMNDPQFGSCETITFGKENSITDYLYPEOLKMTVVKLISHRCQDPHYVGSVTT 357
DB 429 VCLPRADLPDPTTECLSGYKHEALSPTYSERLKAHVRLLYSSSCTGHLNARVTD 488
QY 358 KMLCAAD-----PQWKT--DSQCGDSGGPLVCSLOGRMTLTGIYSWGRGALKKQPGVYTR 411
DB 489 NMLCAGTTRSGGQFQANLHDAQCGDSGGLVCLNDGRMTLVGLISWGGCGKDVPGVYTK 548
QY 412 VSHFLPWIRSHTK 424
DB 549 VTNYLMDIRDMNR 561

```

```

RESULT 8
Q8MB1 ID Q8MB1 PRELIMINARY; PRT; 564 AA.
AC Q8MB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 25, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugita M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1;
DR GO; GO:0005576; F:chymotrypsin activity; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008333; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydroxylase; Kringle; Protease;
KW Serine Protease.
SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6DA937C CRC64;

Query Match 34.5%; Score 826; DB 6; Length 564;
Best Local Similarity 35.6%; Pred. No. 1.9e-70;
Matches 180; Conservative 55; Mismatches 153; Indels 118; Gaps 9;

QY 18 SKGSNELHQVP--SNCD---CLANGTCVSNKYFSNIMHCNPKKFGQCHCEID----- 65
DB 73 SSGGPRCHSVFPVQSGEPRCLNGTCSQALYFSDP--VCQCPREGVGRKCEVDTRARCYED 131
QY 66 ----- 65
DB 132 RGIYRGWTSTESGAQCVNWNSSMLALKEYSGRKPANLRLGLGNHNYCNPNDRTPKWC 191
QY 66 -----KSKTCYEGNGHFYRGKASTDTWGRPCLPMSATVLOQTY 104
DB 192 YVERAGKYSSEFCSTPACSSGKKNYCNPNDRRRPWCYVVGGLKELVQECVNHQCADGKPSPEEL 251
QY 105 HAHRSDALQGLKKNYCNPNDRRRPWCYVVGGLKELVQECVNHQCADGKPSPEEL 164

```

```

Ddb      252  TAQSNANMELGKRNHYCNRPDQSKPMCHVLYKRNKLYIETVDVPOCA----- 299
Qy      165  KFOQC-OKTLRPFPKIIIGSEFTTIENQMPAIIY-RRHRGGSVTVYVCGSIIISPCWISA 222
Ddb      300  --TGGLRQDQKOPQFRIKGGIFTDITAHPMQAAIFNNRRSPDERPLCGGIIINSCWLSIA 357
Qy      223  THCFIDVPKKEDYIVYIGRSRLNSNQGMEKEVENLLIKHODYADTLAHNDIATLKIR 282
Ddb      358  AHCFLEFPQOKRVLKGRITVPLVSAEEQIIEVQPIIHERFDEGT--YNDIATLKIR 415
Qy      283  SKGRCAQPSRITQITCLPSPMYNDPQFGTSGEITGPGKENSITDVLYPEOLKMTVVKLISH 342
Ddb      416  STGSGCAQESQAVRLVCLPDASIQLPDWTCELSIGVGHHEEESPVSFQOLKEAHTVLYPS 475
Qy      343  RECCQPHYYSEVYTWMLCAADPOM-----RTDSCQDPSGGPVLCSLOGKMTITGLVSM 396
Ddb      476  SRCTPQQLKRRITVYGMMLCAGDPRSGAGAVMLHDAQGDGGPVLCTMDGHTILGIIISM 535
Qy      397  GRGCALKDKRGVTVTRVSHFLPWIRSH 422
Ddb      536  GLGCGQGDVGVVTVKVVNLGMIQGH 561

```

RESULT 9

ID	08SQ23	PRELIMINARY;	PRT;	562 AA.
AC	08SQ23;			
DT	01-JUN-2002 (TEMBLrel. 21, Created)			
DT	01-JUN-2002 (TEMBLrel. 21, Last sequence update)			
DT	01-OCT-2003 (TEMBLrel. 25, Last annotation update)			
DE	T-plasminogen activator.			
OC	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Enamel Organ;			
RA	ding Y., Xue J., Bartlett J.D.			
RT	"t-plasminogen activator in tooth tissues."			
PL	Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.			
CC	-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.			
EMBL	AF364605; AAM00297.1; -			
DR	HSSP; P00761; IANI.			
DR	GO; GO:0005576; C:extracellular; IEA.			
DR	GO; GO:0004263; F:chymotrypsin activity; IEA.			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	GO; GO:0004295; F:trypsin activity; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro: IPR009003; Cys_Ser trypsin.			
DR	InterPro: IPR006209; EGF like.			
DR	InterPro: IPR000083; Fibronctn.			
DR	InterPro: IPR006210; IEGF.			
DR	InterPro: IPR000001; Kringle.			
DR	InterPro: IPR001254; Peptidase_S1.			
DR	InterPro: IPR001314; Peptidase_S1A.			
DR	Pfam; PF00008; EGF; 1.			
DR	Pfam; PF00039; fn1; 1.			
DR	Pfam; PF00051; kringle; 2.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PRO0722; CHYMOTRYPSIN.			
DR	PRINTS; PRO0018; KRINGLE.			
DR	ProDom; PD000395; Kringle; 2.			
DR	SMART; SM00181; EGF; 1.			
DR	SMART; SM00058; FN1; 1.			
DR	SMART; SM00130; KR; 2.			
DR	SMART; SM00020; TRYD_Spc; 1.			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 2.			
DR	PROSITE; PS00070; KRINGLE_2; 2.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			

DR PROSITE; PS00134; TRYPSIN_HIS_1.
DR PROSITE; PS00135; TRYPSIN_SER_1.
KM EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KM Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; P9E6B4C77CB101B8 CRC64;

Query Match	34.4%	Score 822.5;	DB 6;	Length 562;
Best Local Similarity	35.2%	Pred. No. 4.2e-70;		
Matches 173; Conservative	61;	Mismatches 168;	Indels 89;	Gaps 8;

QY 20 GSNELHQP-SNCP---CLNGCTVSNKYENHIMCNCPPKKGQOHCEIDSKICYEANG 75
Db 74 GQTOCHSPVYSCSCEPRFCNGCTCQAIYFDF-VCCCPVFATGCEHIDRAICYEDGG 132
QY 76 HFPRKASTDMGRDCLPWSATVYLQOTYAHRSDALQLGKINYYCNPENRRPWCYV 135
Db 133 IYTRWTSTTESGAEVNMVNTTSGLASMPYNRRPDAYLGLGNHNYCNPDKSKPWCYI 192
QY 136 QVGKLPVLOECMVHDCADGKKSPSPBEHLKQ----- 167
Db 193 FKAEEYSPDFSTPACTKEKECECTGKGLDVRGRSLTMSGAPCLPWNLSVLMGKITYAM 252

20

Db	253	NSNAQTGLGHNVCNRPDGTQWCHVLDKHKLWEXYCOLPQCVTGLRQYKRPORIK	312
Qy	181	GGEFTTINQNPWPAIY-RRHGGSVTVYCGSLISPCWVISAITHCFIDYPRKEDYLVLT	239
Db	313	GGVLADITSHPMQAAIFVKNRRSPGEEFRLCGGILISSCWVLSAHCFOEFPFPHVAVLL	372
Qy	240	GRSLNSNTQGEEMFEVEYNLIHHDYASADTLAHNDI.ALKIKRSEKRCAPSPRTIOTIC	299
Db	373	GRTRILYPGEEBQAFVEYKTVHKEFDDT--YNDIALALQLSDSLSLCAQESDAKIVC	430
Qy	300	LPSMYNDPQFQTSCEITGFGKENSITDYLYPEQLMTYVKLIISHRECOOPHYGSEVTTKM	359
Db	431	LPEANLQLPDWTECELISGYGKHGSAISPFYSRLKEAHVRLYPPSSRCTSKLFNKTIINNM	490
Qy	360	LCADPQW-----KTSQCGDSGGLPYCLISQSMITLTVISVNRGCLADKPYVTRVS	413
Db	491	LCADDTSGSGGNALHDAICGDSGGPLVCMKGNHMLVGVISWGLGCGQXDPVPGVTKYT	550
Qy	414	HFLPWIRSHTK 424	
Db	551	NYLNMIRDNTR 561	

RESULT 10

Q9BZW1	PRELIMINARY;	PRT;	395 AA.
AC	Q9BZW1		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Neonatal chromobolytic agent alpha-form (Fragment).		
OS	Homo sapiens (Human).		
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
EN	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Dou D.;		
RT	"A brain-type plasminogen activator.";		
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: BELONGS TO THE EMBL/GENBANK/DBJ DATABASES.		
CC	-1- SIMILARITY: CONTAINS 1 KININASE DOMAIN.		
DR	EMBL; AF260825; AKK1956.1; -.		
DR	HSSP; P00750; 1PK2.		
DR	GO; GO:0005576; C:extracellular; IEA.		
DR	GO; GO:0004263; F:chymotrypsin activity; IEA.		
DR	GO; GO:0008233; F:peptidase activity; IEA.		
DR	GO; GO:0004295; F:trypsin activity; IEA.		
DR	GO; GO:0006508; P:trypsin activity; IEA.		

DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR000083; Fibronctn.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00039; fn1.1.
 DR Pfam: PF00051; kringle.1.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR0018; KRINGLE.
 DR ProDom: PD000395; Kringle.1.
 DR SMART: SM00058; FN1.1.
 DR SMART: SM00130; KR.1.
 DR SMART: SM00020; tryp_Spc.1.
 DR PROSITE: PS01253; FIBRONECTIN_1.1.
 DR PROSITE: PS00021; KRINGLE_1.1.
 DR PROSITE: PS00070; KRINGLE_2.1.
 DR PROSITE: PS0240; TRYPSIN_DOM.1.
 DR PROSITE: PS00134; TRYPSIN_HIS.1.
 DR PROSITE: PS00135; TRYPSIN_SER.1.
 DR GlycoProtex: Hydrolase; Kringle; Protease; Serine protease.
 FT NON_TER 395
 SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 32.8%; Score 784.5; DB 4; Length 395;
 Best Local Similarity 42.0%; Pred. No. 1.2e-66;
 Matches 161; Conservative 49; Mismatches 144; Indels 29; Gaps 8;

QY 50 WCNCPPKFGGQHGEIDSKTCYGNHGFYRGKASTDTMGRPCLPMSATVLOQTTHAHS 109
 DB 33 WCNS---GAQGS--EGNSDCYFGNSAYNGTSHLSTESGASCLPMNSMILIGVYTAQNP 87
 QY 110 DALQLGKHNKYNCRPNDRRRPWCYVQVGLKPLVQECMVNDGADGKKSPPELKFQCG 169
 DB 88 SAQALGKHNKYNCRPNDRRRPWCYVQVGLKPLVQECMVNDGADGKKSPPELKFQCG 169
 QY 170 -QKTLRPRKIIIGSEFTTLENQWPAIYRRH-RGSSVTVYCGSLISPCWISATHCFI 227
 DB 134 LRQYQPOFQIRKIGLFPADISHPMQAAIFAKHEESPERELCGIILSSCMILSAHCFQ 193
 QY 228 DYRKEDYIVYGRSLNSTQGEKFEVENLILHSDYADTLAHNDILALKIRSEGR 287
 DB 194 ERPPHHLTVILGRYRVVPGEEBQKFECEKTYVHKEFDDT--YNDIALQLKSDSSR 251
 QY 288 CAQSRKTIQICLPMTNDPQFGTSCETIFGKENSSTDYLYPQKMTVVKLISHBECQ 347
 DB 252 CAQSESVKRVCLPPADLPQDPTBCELSIGYKHEALSPYSERLKAHRLYPSRCTS 311
 QY 348 PHYGSEVTTKMLCAAD-----PQWKT-DSQGDSSGGLVCSLOGRWTLTGIVSWGRC 401
 DB 312 QHLNRTVTDNMLCAGDTRSGQFQAMLHACQDSSGGLVCLNDRWTLVGIIISWGLGCG 371
 QY 402 LKDKRGVYTRVSHFLPMISHTK 424
 DB 372 QKQVPGVYTKVTNYLDIMRDMR 394

RESULT 11

Q800Y7 PRELIMINARY; PRT; 540 AA.
 AC Q800Y7;
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Hepatocyte growth factor activator (Fragment).
 OS Melospiza gallopavo (Common turkey).
 OC Archosauria; Melosauria; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Melagris.
 OX NCBI_TaxID=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=2212796; PubMed=12128063;
 RA Holsberger D.R., Becker A.E., Thureston R.J., Rice C.D.;

RT "Expression of a hepatocyte growth-factor activator protein in turkey
 (Melagris gallopavo) deferent duct epithelial cells."
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AY216598; AA046038.1; -
 DR GO; GO:0005576; Cytoplasmic; IEA.
 DR GO; GO:0004263; F-actin polymerization; IEA.
 DR GO; GO:0004295; F-actin polymerization; IEA.
 DR GO; GO:0005508; P-actin polymerization; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR006209; EGF-like.
 DR InterPro: IPR000683; Fibronctn.
 DR InterPro: IPR000562; FN_Type_II.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00051; kringle.1.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR0018; KRINGLE.
 DR ProDom: PD000395; Kringle.1.
 DR SMART: SM00058; FN1.1.
 DR SMART: SM00059; FN2.1.
 DR SMART: SM00130; KR.1.
 DR SMART: SM00020; tryp_Spc.1.
 DR PROSITE: PS00022; EGF_2.1.
 DR PROSITE: PS01186; EGF_2.1.
 DR PROSITE: PS01253; FIBRONECTIN_1.1.
 DR PROSITE: PS00021; KRINGLE_1.1.
 DR PROSITE: PS00070; KRINGLE_2.1.
 DR PROSITE: PS0240; TRYPSIN_DOM.1.
 DR PROSITE: PS00134; TRYPSIN_HIS.1.
 DR PROSITE: PS00135; TRYPSIN_SER.1.
 FT NON_TER 1
 SQ SEQUENCE 540 AA; 61159 MW; 0BB3BA489C0B577F CRC64;

Query Match 30.6%; Score 731.5; DB 13; Length 540;
 Best Local Similarity 37.2%; Pred. No. 2.2e-61;
 Matches 159; Conservative 67; Mismatches 166; Indels 35; Gaps 11;

QY 11 CVLVSDSKGNSNHLQVPSNCDCLNGTGC-----VSNKYSNTHMCNCPKPKFGGQHGEIDK 66
 DB 121 CTVDSQDMCLNVEHKSCHDPCNMGCECKMIASSGKT-----CDCKGPFVGYKCNIVP 175
 QY 67 SKTCYGNHGFYRGKASTDTMGRPCLPMSATVLOQTTHAHS DALQLGKHNKYNCRPN 126
 DB 176 NHHCYRNGCTEYRGTAKTITISGHSCLPMNSDLYRLHVDSEYKAAVQLGIFGFSYCRNDP 235
 QY 127 NRRBPWCYVQVGLKPLVQECMVNDGADGKKSPPEL-----KFGCGK-----TLRP 175
 DB 236 EDEKPCYITMKDONSLSWEYCNITSCASRRRPVLEDTFAPVPRPCGRHKGKSFVRP 295
 QY 176 RFKIIIGSEFTTLENQWPAIYRRHRSQSVTVYCGSLISPCWISATHCFIDYPKEDY 235
 DB 296 R--IIIGSSSLGSHWTAIY---IGES---FAGTLIQTCMVVSAHCFANSPPKSSI 347
 QY 236 IYLLGRSLNSTQGEKFEVENLILHSDYADTLAHNDILALKIRSEGRCAQDSRTI 295
 DB 348 KYVLGHPNRTYDTQTEIEIKYILYQYSVFRPTEN-DIALIKLKNGGRCAYVKSQEV 406
 QY 296 QTCICLP---SMNDPQFGTSCETIFGKENSSTDYLYPQKMTVVKLISHBECQPHYNG 352
 DB 407 QPCLPESNTVFPD-QF--KQIISGKHKHENTIGSDVLOQTTLPIIEBKCRSEFETIG 463
 QY 353 SEVTTKMLCAADPQWKTDSQGDSSGGLVCSLOGRWTLTGIVSWGRCALKDKPGVYTRV 412

```
Db 464 TEISENNFCCGYRDSKSDACGDSGGFLACENNEISITLYGVISWGDCGAVNRPGYTRY 523
Qy 413 SHFLPWT 419
Db 524 PNYVNW 530

RESULT 12
Q8VCS4 PRELIMINARY; PRT; 653 AA.
ID 08VCS4
AC 08VCS4;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strauberg R.;
CC Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1 SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; IAN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTYPET1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174EDBFDD CRC64;
```

```
Qy 11 CYLVVSDSKSNELHQPENSCDLCNGTCVSNKFFSNIM-----CNCPEKFGGHC 62
Db 225 CGMGEGARCEPTHTHTACLSPLCNGTC-----HLVGTGTSVTCGLGYAGRC 275
Qy 63 EIDSKTCYEGNGHFRGASTDTWGRPCLPWN SATVLQOTYHAHSDALDGLGKANYC 122
Db 276 NIVPTEHCFLGNOTEXRVAVASTAGSLCLANMSDLVDELHVDSVAALVLLGPHAAVC 335
Qy 123 RNPDRRRPWCYVQVQLKVLQECVWHDCA- GKRPSPPEEL-----KFGCGOK 171
Db 336 RNPDRKRPWCYVQVQKNSLWSEYCRLTACE SARVHSGSPETLALPESAPVRPFCGR 395
Qy 172 -----TLRPRFKLIGEFITINQFWFAIYRRHRGSGTYVYCGSLISPCWISATNCF 226
Db 396 HKKRTLRPR--IIGSSSLPGSHPLAIIY---IGNS---FCAGSLVHTCWVVAACHF 447
Qy 227 IDVPKKEDYIVYIGSRINSNTQGENKFVEVNLILHKDYSADTLAHNDIALIKIRSKG 286
Db 448 ANSPRPDSITVVLGQHFNRRTDVTQTFGIEKVPYTLVSFNPNNH-DLVILRLKKGE 506
Qy 287 RCAQPSRTIQTCLPSMNDPQRTSCETTFGKENSITDYLPEQLKMTVVLISHREG 346
Db 507 RCARNSQFVQPICLPAGSGPFGHKCQLAGWMDENVSSYNSLLEALVPLVADHKCS 566
Qy 347 QPHYSGEYVTKMLCAADPQMKTSCCGDSGGFLVCSLGRMTLTGIVSMGRGALKDXP 406
Db 567 SPFVYQADISPNMLCAGYFDCKSDACGDSGGFLVEKRGVAVLYGIIEMGCGGRLNKP 626
Qy 407 GYTVRVSHFLPWIRSHTK 424
Db 627 GYTVRVANVYDWINDRIR 644

RESULT 13
ID 097507 PRELIMINARY; PRT; 616 AA.
AC 097507;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Khara T.;
RT "Porcine liver factor XII."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1 SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
```

DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; ENTYREIT.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin DOM; 1.
 DR PROSITE; PS00134; TRYPsin HIS; 1.
 DR PROSITE; PS00135; TRYPsin SER; 1.
 KM EGF-like domain; Glycoprotein; Hydrolyase; Kringle; Protease;
 KM Serine protease.
 SQ SEQUENCE 616 AA; 66012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 29.0%; Score 693.5; DB 6; Length 616;
 Best Local Similarity 35.3%; Pred. No. 1,1e-57;
 Matches 161; Conservative 61; Mismatches 159; Indels 75; Gaps 14;

QY QVNSNCDCLNGGTGVSNKYPFSNIHMCNCPKFGGHCIDSKTYEGNGHFRGKASTD 85
 DB 176 QVSTNPDCLNGGSCLOTE---GHRLCGCPGVAGRLCDVDLKRQCVSDRGLSRGMAQT 232
 QY 86 TMRPCLPMSATVLOQTY-HARRSDALQGLGKHNVCNPDNRKRRPWCYVQGLKPLVQ 144
 DB 233 LSGAPCPMNAS---EATYNNMTAEQALNMGLDHAFCRNPDNDTRPWCYVQGLKPLVQ 288
 QY 145 EGVVHPCAD--GKKP-----SSPPEL- 164
 DB 289 YCLARCOAIGERFPLITLTQSPSEKQDPLSRPQPTTQPSQULTAMKAPRQRG 348
 QY 165 -----KFGCGKTLRPRF---KIIGSEPTTIENQWFAIYRRHGGSVTVYCGSLI 214
 DB 349 PLPSAGLVGGQR-LRRRLSLNRIVGVALPAGAHYIALYWGQ-----FCAGSLI 401
 QY 215 SPQWVLSATCFIDYKEDYIYLGSRNLNSTQGMKEVENLIHKQYSDTLAHNN 274
 DB 402 APCMVLTAAHCLQNRPAPELTVLQGDHRNOSCEQCOTLAIVRSYRLHESYSPKTYQH-- 459
 QY 275 DIALLKIR-SKEGRCAQSPRTIQTICP---SMYNDQFQTSCEITGFGKENSVDIYPE 330
 DB 460 DIALVRKETADSCAPSPFPVQVCLPRSVASAEPE-CALCEVAMGQFQCAEETSS 518
 QY 331 QLMTVTKLISHRECOQPHYYSVTTKMLCAADPWKTKSCQDSSGGLPVC---SLQGR 387
 DB 519 FLDEAGVPLSPRCSADVDHGAFTPGMLCAGTLEGDTACGDSGGLPVCDETERAQ 578
 QY 388 MTLTGIVSGRGALCKDKPGVYTRVSHFLPWRSH 423
 DB 579 LVLRGIVSWGSGGDRLPQVYTDVANYLAMIQHT 614

RESULT 14

ID 081225 PRELIMINARY; PRT; 615 AA.
 AC 081225;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Coagulation factor XII-Mie.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wada H., Nishio K., Nakatani K., Kasei Y., Abe Y., Nobori T.,
 RT "Molecular characterization of coagulation factor XII-Mie."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBD databases.
 DR EMBL; AB095845; BAC23095.1; -
 DR GO; GO:0005576; Cextracellular; IEA.
 DR GO; GO:0005509; Fcalcium ion binding; IEA.
 DR GO; GO:0004293; Fchymotrypsin activity; IEA.
 DR GO; GO:0004295; Ftrypsin activity; IEA.
 DR GO; GO:0006508; Pproteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser trypsin.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000083; Fibrinectn.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; FN1; 1.
 DR Pfam; PF00040; FN2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; ENTYREIT.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00179; EGF_CA; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin DOM; 1.
 DR PROSITE; PS00134; TRYPsin HIS; 1.
 DR PROSITE; PS00135; TRYPsin SER; 1.
 SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 28.8%; Score 689; DB 4; Length 615;
 Best Local Similarity 34.3%; Pred. No. 3,1e-57;
 Matches 153; Conservative 59; Mismatches 164; Indels 70; Gaps 9;

QY 33 CLNGTGVSNKYPFSNIHMCNCPKFGGHCIDSKTYEGNGHFRGKASTDTMRPCL 92
 DB 183 CLHGRCLE---VEGRLCHCPVGYGPRCDVTKASCYDRGLSYRLARTLISGAPQ 239
 QY 93 PMSATVLOQTY-HARRSDALQGLGKHNVCNPDNRKRRPWCYVQGLKPLVQECVHQC 151
 DB 240 PMAS---EATYNNMTAEQALNMGLDHAFCRNPDNDTRPWCYVQGLKPLVQECVHQC 295
 QY 152 -----ADGKKSPPEL- 164
 DB 296 QPTQAAPRPVPSRLHVLMPAQPAPKRPQPTTRTPPOSQTPGALPAKRGQPSLVRNG 355
 QY 165 KFGCGG---KTLRPRKIIGSEPTTIENQWFAIYRRHGGSVTVYCGSLISPQWIS 221
 DB 356 PLSCGRLKRSLSMSRIVGVALPAGAHYIALYWGHS-----FCAGSLIAPCVLIT 409
 QY 222 ATTCFIDYKEDYIYLGSRNLNSTQGMKEVENLIHKQYSDTLAHNDILAKI 281
 DB 410 AAHCLQDRAPBELTVLQGDHRNNSCEPCQTLAIVRSYRLHESYSPKTYQH--PVSQYHDLALRL 467

```

QY 282 R-SKEGRCAQPSRTIQTICLPSMWNIDPQFTSCETISFGKENSTDYLYPEQLKMTVVKLI 340
DB 468 QENADSGCALLSPVQVCLPSCAARSSETLLCQVACGCGHFGAEYSAFLGEAVPFL 527
QY 341 SHRECCQPHYYSSEVTTKMLCAADPQWKTDSCQDSSGGLVCSLQ--RWTLTGIVSWG 397
DB 528 SLRCSAPVDHSGSILPQMLCAGFLGEGTDACQDSSGGLVCEPQAEERLTLQGIISWG 587
QY 398 RGCALDKRGVYTRVSHFLPWIRSH 423
DB 588 SGGCDNRKPGVYTDVAYVLYAMIREHT 613

RESULT 15
ID 035727 PRELIMINARY; PRT; 597 AA.
AC 035727;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Schloesser M., Schwager S., Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X9571; CA67891.1; -.
DR HSSP; P00760; 1A07.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinectin.
DR InterPro; IPR000562; FN_type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00068; EGF; 2.
DR Pfam; PF00039; FN1; 1.
DR Pfam; PF00040; FN2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.

```

```

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine Protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 28.4%; Score 680.5; DB 11; Length 597;
Best local similarity 34.7%; Pred. No. 1,9e-56;
Matches 156; Conservative 65; Mismatches 169; Indels 59; Gaps 12;

QY 17 DSKGSELEH-----QVPSNCDCINGTCVSNKYFSNIMHCNCPKKFGQAHCEIDKSKTC 70
DB 162 ECKGS-EAHCKPVASQACINPLNGSCL---LVEDHPLCRCPGTGTGTFCDLDMATC 217
QY 71 YEGNGHFYRKASTDWGRPCLPMNSATVLAQTY-HAHSDALQGLGKANTCRNDNR 129
DB 218 YERGLSTYRGAGTTSQAPCQRM---TVEATYRMTEKQALSMGLGHNAFCRNDNDT 273
QY 130 RPYCYQVGLKPLVQECMVHDC-----ADGKRPSSBP----- 161
DB 274 RPYCFWMSGDRLSMDYCGLEOCQTPYFAPLVVPESSQESPQAPSLSHAPNDSTHQSL 333
QY 162 -ELKFGCGQ---KTLRPFKIIIGBEFTTENQPFALYRRHRGGSVTVYVCGSLISPC 217
DB 334 SKTNMGCCQRFERKGLSSFMRVVGGVLALPGSHPIALYKNN-----FCAGSLIATC 387
QY 218 WVISATHCFTDYPKEDYIVYIGSRILNSNTQEMKFEVENTLHKDYADTLAHHNDIA 277
DB 388 WVTIAHCLQNRAPBELTVVIGQDRHNSCEMCQTLAVRSYRLHGFSSITYQH--DLA 445
QY 278 LKIR-SKEGRCAQPSRTIQTICLPSMWNIDPQFTSCETISFGKENSTDYLYPEQLKMTV 336
DB 446 LRLQESKTNSCAILSPHQVPCLPSGAAPPSRTVCEVAGWQHOLEGAEEYSTFLQEAQ 505
QY 337 VKLSHRECCQPHYYSSEVTTKMLCAADPQWKTDSCQDSSGGLVCSLQ--RWTLTGIVSWG 393
DB 506 VPFIALDRCSNSVHGDALIPGMLCAGFLGEGTDACQDSSGGLVCEBGTAEHQTLRGV 565
QY 394 VSWGRCALDKRGVYTRVSHFLPWIRSH 422
DB 566 ISWGSQGNRKNKPGVYTDVANYLYAMIREHT 613

```

Search completed: March 18, 2004, 11:19:17
 Job time : 119.888 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 1.65427 Seconds
(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KKG 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	4	5	AAAG79461
2	22	100.0	7	4	AAAG4844
3	22	100.0	7	4	AAAG4233
4	22	100.0	7	4	AAAG4849
5	22	100.0	7	4	AAAG3769
6	22	100.0	7	4	AAAG45363
7	22	100.0	7	4	AAAG45898
8	22	100.0	7	4	AAAG45884
9	22	100.0	7	4	AAAG45903
10	22	100.0	7	4	AAAG4641
11	22	100.0	8	4	AAAG35204
12	22	100.0	10	4	AAAG84320
13	22	100.0	10	4	AAAG6341
14	22	100.0	10	4	AAAG6245
15	22	100.0	10	4	AAAG6183
16	22	100.0	10	4	AAAG3019
17	22	100.0	11	2	AAAG16616
18	22	100.0	12	2	AAAG3013
19	22	100.0	12	5	AAAG2420
20	22	100.0	13	2	AAAG30274
21	22	100.0	13	2	AAAG10774
22	22	100.0	13	5	AAAG2487
23	22	100.0	14	2	AAAG13742
24	22	100.0	14	6	AAAG1869
25	22	100.0	15	2	AAAG13741

26	22	100.0	15	2	AAAG6300	AAAG6300	Magainin
27	22	100.0	15	2	AAAG2245	AAAG2245	Magainin
28	22	100.0	16	2	AAAG8109	AAAG8109	Alpha-hel
29	22	100.0	17	2	AAAG9102	AAAG9102	Magainin
30	22	100.0	17	2	AAAG2589	AAAG2589	Truncated
31	22	100.0	17	6	AAAG1913	AAAG1913	Human int
32	22	100.0	18	2	AAAG10659	AAAG10659	Magainin
33	22	100.0	18	2	AAAG10661	AAAG10661	Magainin
34	22	100.0	18	2	AAAG13929	AAAG13929	Calcitonin
35	22	100.0	18	2	AAAG15796	AAAG15796	Magainin
36	22	100.0	18	2	AAAG15798	AAAG15798	Magainin
37	22	100.0	18	2	AAAG9101	AAAG9101	Magainin
38	22	100.0	18	2	AAAG94000	AAAG94000	Truncated
39	22	100.0	18	2	AAAG6317	AAAG6317	Magainin
40	22	100.0	18	2	AAAG6319	AAAG6319	Magainin
41	22	100.0	18	2	AAAG22250	AAAG22250	Magainin
42	22	100.0	18	2	AAAG22248	AAAG22248	Magainin
43	22	100.0	19	2	AAAG9100	AAAG9100	Magainin
44	22	100.0	19	2	AAAG3999	AAAG3999	Truncated
45	22	100.0	19	3	AAAG7472	AAAG7472	Magainin

ALIGNMENTS

RESULT 1
AAAG79461
ID AAAG79461 standard; peptide; 4 AA.
AC
XX
DT 15-NOV-2002 (first entry)
XX
DE sec-uPA long A chain peptide fragment 1, amino acids 55-58.
XX
XX
KM Single chain prepro-urokinase; sc-uPA, pro-urokinase; HIV;
KM high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KM long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KM low molecular weight urokinase-type plasminogen activator; LMW-uPA; CD87;
KM binding domain.
OS Homo sapiens.
XX
PN EP1232755-A2.
XX
PD 21-AUG-2002.
XX
PF 15-FEB-2002; 2002EP-00003555.
XX
PR 20-FEB-2001; 2001JP-00042655.
PR 19-JUN-2001; 2001JP-00184284.
XX
PA (JCRP-) JCR PHARM CO LTD.
XX
PI Wada M, Wada N;
XX
DR WPI; 2002-610512/66.
XX
PT Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g. high
PT molecular weight urokinase-type plasminogen activator, amino-terminal
PT fragment or an anti-CD87 antibody.
XX
XX
PS Example; Page 23; 38pp; English.
XX
CC The sequences given in AAAG79461-63 represent peptide fragments derived
CC from the long A chain of single chain prepro-urokinase (sc-uPA). Pro-
CC urokinase (amino acids 21-431) with a cleavage between amino acids 178
CC and 179 gives high molecular weight urokinase-type plasminogen activator
CC (HMW-uPA). HMW-uPA is a protein consisting of two peptide chains linked
CC by a di-sulphide bond. The chains, long A and B, are formed by enzymatic
CC cleavage between amino acids 178 and 179 of pro-urokinase. HMW-uPA
CC includes an EGF-like domain, a kringle domain and a urokinase receptor
CC (CD87) binding domain. HMW-uPA is then cleaved between amino acids 155

CC and 156 to give low molecular weight urokinase-type plasminogen activator
 CC (LMM-uPA) (amino acids 156-178 and 179-431), that has no plasminogen
 CC activator activity. sc-uPA, or these fragments of it, may be used in the
 CC anti-HIV agents of the invention which comprise a ligand molecule that
 CC binds to CD87. The agents are useful for treating HIV-infected humans for
 CC suppression of reproduction of HIV. The anti-HIV agents act by a
 CC mechanism of action different from those of conventional drugs, widening
 CC the choice of therapeutics agents and avoiding problems of resistant HIV
 CC XX

SO Sequence 4 AA:

Query Match 100.0%; Score 22; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 DB 1 KKEG 4

RESULT 2

AA44844 standard; peptide; 7 AA.

AA44844;

25-OCT-2001 (first entry)

H11 binding site consensus conforming peptide (CCP) #1115.

Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
 immunogenically cross-reactive; cancer; immunogenic cancer cell;
 cytostatic; vaccine; tumour-specific immunogenic response inducer;
 astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
 ependymoma; medulloblastoma; primitive neural ectodermal tumour.

Homo sapiens.
 Synthetic.

CA2290722-A1.

08-JUN-2001.

08-DEC-1999; 99CA-02290722.

08-DEC-1999; 99CA-02290722.

(NOVO-) NOVOPHARM BIOTECH INC.

Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
 Entwistle JM, Macdonald GC;

WPI; 2001-425937/46.

Composition useful for treating and diagnosing cancer, comprises stress
 protein-peptide complexes associated with tumor, and isolated antigen-
 binding fragments of an antibody that binds specifically to the complex.

Example 4; Page 105; 154pp; English.

The present invention describes a composition (I) comprising stress
 protein-peptide complexes (SPPC) associated with tumors that is
 specifically immunogenically cross-reactive with cell surface-associated
 SPPCs specific to target cancer (TC). Also described is an isolated
 antigen-binding fragment of an antibody that binds specifically to SPPCs
 or a population of different SPPCs consisting of immunogenic cancer cell
 surface-associated SPPC of TC. (I) has cytostatic activity and can be
 used in vaccine production and as a tumour-specific immunogenic response
 inducer. (I) is useful for treating 71 types of cancers or tumors in a
 subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
 oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 ectodermal tumour (PNET). (I) is useful as cancer immunogen including
 vaccines. (I) is useful for diagnostic and palliative use, for detecting

CC or imaging cancer cells, and to monitor the course of amelioration of
 CC malignancy in an individual. AA43707 to AA47109 represent peptides
 CC which are used in the exemplification of the present invention
 CC XX

SO Sequence 7 AA:

Query Match 100.0%; Score 22; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 DB 2 KKEG 5

RESULT 3

AA44233 standard; peptide; 7 AA.

AA44233;

25-OCT-2001 (first entry)

H11 binding site consensus conforming peptide (CCP) #504.

Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
 immunogenically cross-reactive; cancer; immunogenic cancer cell;
 cytostatic; vaccine; tumour-specific immunogenic response inducer;
 astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
 ependymoma; medulloblastoma; primitive neural ectodermal tumour.

Homo sapiens.
 Synthetic.

CA2290722-A1.

08-JUN-2001.

08-DEC-1999; 99CA-02290722.

08-DEC-1999; 99CA-02290722.

(NOVO-) NOVOPHARM BIOTECH INC.

Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
 Entwistle JM, Macdonald GC;

WPI; 2001-425937/46.

Composition useful for treating and diagnosing cancer, comprises stress
 protein-peptide complexes associated with tumor, and isolated antigen-
 binding fragments of an antibody that binds specifically to the complex.

Example 4; Page 102; 154pp; English.

The present invention describes a composition (I) comprising stress
 protein-peptide complexes (SPPC) associated with tumors that is
 specifically immunogenically cross-reactive with cell surface-associated
 SPPCs specific to target cancer (TC). Also described is an isolated
 antigen-binding fragment of an antibody that binds specifically to SPPCs
 or a population of different SPPCs consisting of immunogenic cancer cell
 surface-associated SPPC of TC. (I) has cytostatic activity and can be
 used in vaccine production and as a tumour-specific immunogenic response
 inducer. (I) is useful for treating 71 types of cancers or tumors in a
 subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
 oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 ectodermal tumour (PNET). (I) is useful as cancer immunogen including
 vaccines. (I) is useful for diagnostic and palliative use, for detecting
 or imaging cancer cells, and to monitor the course of amelioration of
 malignancy in an individual. AA43707 to AA47109 represent peptides
 CC which are used in the exemplification of the present invention
 CC XX

SO Sequence 7 AA:

Query Match 100.0%; Score 22; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRF 4
Db 2 KRF 5

RESULT 4

AAM44849
ID AAM44849 standard; peptide; 7 AA.

AC AAM44849;

DT 25-OCT-2001 (first entry)

DE H11 binding site consensus conforming peptide (CCP) #1120.

KM Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;

KM immunogenically cross-reactive; cancer; immunogenic cancer cell;

KM cytostatic; vaccine; tumour-specific immunogenic response inducer;

KM astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;

KM ependymoma; medulloblastoma; primitive neural ectodermal tumour.

OS Homo sapiens.

OS Synthetic.

PN CA2290722-A1.

PD 08-JUN-2001.

PF 08-DEC-1999; 99CA-02290722.

PR 08-DEC-1999; 99CA-02290722.

XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

XX Entwistle JM, MacDonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress

XX protein-peptide complexes associated with tumour, and isolated antigen-

XX binding fragments of an antibody that binds specifically to the complex.

XX Example 4; Page 105; 154pp; English.

XX The present invention describes a composition (I) comprising stress

XX protein-peptide complexes (SPPC) associated with tumours that is

XX specifically immunogenically cross-reactive with cell surface-associated

XX SPPCs specific to target cancer (TC). Also described is an isolated

XX antigen-binding fragment of an antibody that binds specifically to SPPCs

XX or a population of different SPPCs consisting of immunogenic cancer cell

XX surface-associated SPPC of TC. (I) has cytostatic activity and can be

XX used in vaccine production and as a tumour-specific immunogenic response

XX inducer. (I) is useful for treating 71 types of cancers or tumours in a

XX subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,

XX oligodendroglioma, ependymoma, medulloblastoma, and primitive neural

XX ectodermal tumour (PNET). (I) is useful as cancer immunogen including

XX or imaging cancer cells, and to monitor the course of amelioration of

XX malignancy in an individual. AAM43707 to AAM47109 represent peptides

XX which are used in the exemplification of the present invention

XX Sequence 7 AA;

XX Query Match 100.0%; Score 22; DB 4; Length 7;

XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRF 4
Db 2 KRF 5

RESULT 5
AAM43769
ID AAM43769 standard; peptide; 7 AA.

AC AAM43769;

DT 25-OCT-2001 (first entry)

DE H11 binding site consensus conforming peptide (CCP) #40.

KM Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;

KM immunogenically cross-reactive; cancer; immunogenic cancer cell;

KM cytostatic; vaccine; tumour-specific immunogenic response inducer;

KM astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;

KM ependymoma; medulloblastoma; primitive neural ectodermal tumour.

OS Homo sapiens.

OS Synthetic.

PN CA2290722-A1.

PD 08-JUN-2001.

PF 08-DEC-1999; 99CA-02290722.

PR 08-DEC-1999; 99CA-02290722.

XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

XX Entwistle JM, MacDonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress

XX protein-peptide complexes associated with tumour, and isolated antigen-

XX binding fragments of an antibody that binds specifically to the complex.

XX Example 4; Page 101; 154pp; English.

XX The present invention describes a composition (I) comprising stress

XX protein-peptide complexes (SPPC) associated with tumours that is

XX specifically immunogenically cross-reactive with cell surface-associated

XX SPPCs specific to target cancer (TC). Also described is an isolated

XX antigen-binding fragment of an antibody that binds specifically to SPPCs

XX or a population of different SPPCs consisting of immunogenic cancer cell

XX surface-associated SPPC of TC. (I) has cytostatic activity and can be

XX used in vaccine production and as a tumour-specific immunogenic response

XX inducer. (I) is useful for treating 71 types of cancers or tumours in a

XX subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,

XX oligodendroglioma, ependymoma, medulloblastoma, and primitive neural

XX ectodermal tumour (PNET). (I) is useful as cancer immunogen including

XX or imaging cancer cells, and to monitor the course of amelioration of

XX malignancy in an individual. AAM43707 to AAM47109 represent peptides

XX which are used in the exemplification of the present invention

XX Sequence 7 AA;

XX Query Match 100.0%; Score 22; DB 4; Length 7;

XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6	
AA45363	
ID	AA45363 standard; peptide; 7 AA.
XX	
AC	AA45363;
XX	
DT	25-OCT-2001 (first entry)
XX	
DE	H11 binding site consensus conforming peptide (CCP) #1634.
XX	
KW	Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW	immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW	cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW	astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;
KW	ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX	
OS	Homo sapiens.
OS	Synthetic.
PN	CA2290722-A1.
PD	08-JUN-2001.
PF	08-DEC-1999; 99CA-02290722.
PR	08-DEC-1999; 99CA-02290722.
PA	(NOVO-) NOVOPHARM BIOTECH INC.
PI	Kaplan HA, Maifi PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI	Entwistle JM, Macdonald GC;
DR	WPI: 2001-425937/46.
PT	Composition useful for treating and diagnosing cancer, comprises stress
PT	protein-peptide complexes associated with tumor, and isolated antigen-
PR	binding fragments of an antibody that binds specifically to the complex.
XX	
PS	Example 4; Page 106; 154pp; English.
XX	
CC	The present invention describes a composition (I) comprising stress
CC	protein-peptide complexes (SPPC) associated with tumors that is
CC	specifically immunogenically cross-reactive with cell surface-associated
CC	SPPCs specific to target cancer (TC). Also described is an isolated
CC	antigen-binding fragment of an antibody that binds specifically to SPPCs
CC	or a population of different SPPCs consisting of immunogenic cancer cell
CC	surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC	used in vaccine production and as a tumour-specific immunogenic response
CC	inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC	subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,
CC	oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC	ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC	vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC	or imaging cancer cells, and to monitor the course of amelioration of
CC	malignancy in an individual. AA43707 to AA47109 represent peptides
CC	which are used in the exemplification of the present invention
XX	
SO	Sequence 7 AA;
QY	
DB	
Query Match	100.0%; Score 22; DB 4; Length 7;
Best Local Similarity	100.0%; Pred. No. 1.4e+06;
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KKEG 4	
2 KKEG 5	
RESULT 7	
AA45898	
ID	AA45898 standard; peptide; 7 AA.
XX	
AC	AA45898;

XX	25-OCT-2001	(first entry)
DT		
XX		
XX	H11 binding site consensus conforming peptide (CCP) #2169.	
XX		
XX	Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPpC;	
KW	immunogenically cross-reactive; cancer; immunogenic cancer cell;	
KM	cytostatic; vaccine; tumour-specific immunogenic response inducer;	
KM	astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;	
KM	ependymoma; medulloblastoma; primitive neural ectodermal tumour.	
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
PN	CA2290722-A1.	
XX		
PD	08-JUN-2001.	
XX		
PF	08-DEC-1999; 99CA-02290722.	
XX		
PR	08-DEC-1999; 99CA-02290722.	
XX		
PA	(NOVO-) NOVOPHAM BIOTECH INC.	
XX		
XX	Kaplan HA, Maiti PK, Faat DG, Herman W, Dan MD, Lewis KE;	
PI	Entwistle JM, Macdonald GC;	
XX		
DR	WPI; 2001-425937/46.	
XX		
PT	Composition useful for treating and diagnosing cancer, comprises stress	
PT	protein-peptide complexes associated with tumor, and isolated antigen-	
PT	binding fragments of an antibody that binds specifically to the complex.	
XX		
PS	Example 4; Page 108; 154pp; English.	
XX		
CC	The present invention describes a composition (I) comprising stress	
CC	protein-peptide complexes (SPpC) associated with tumors that is	
CC	specifically immunogenically cross-reactive with cell surface-associated	
CC	SPpCs specific to target cancer (TC). Also described is an isolated	
CC	antigen-binding fragment of an antibody that binds specifically to SPpCs	
CC	or a population of different SPpCs consisting of immunogenic cancer cell	
CC	surface-associated SPpC of TC. (I) has cytostatic activity and can be	
CC	used in vaccine production and as a tumour-specific immunogenic response	
CC	inducer. (II) is useful for treating 71 types of cancers or tumors in a	
CC	subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,	
CC	oligodendroglioma, ependymoma, medulloblastoma, and primitive neural	
CC	ectodermal tumour (PNET). (I) is useful as cancer immunogen including	
CC	vaccines. (I) is useful for diagnostic and palliative use, for detecting	
CC	or imaging cancer cells, and to monitor the course of amelioration of	
CC	malignancy in an individual. AAM43707 to AAM47109 represent peptides	
CC	which are used in the exemplification of the present invention	
XX		
XX	Sequence 7 AA;	
XX		
QY	1 KKEG 4	
DB	2 KKEG 5	
XX		
XX	100.0%; Score 22; DB 4; Length 7;	
XX	Best Local Similarity 100.0%; Pred. No. 1.4e+06;	
XX	Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX		
XX	RESULT 8	
XX	AAM44584	
XX	AAM44584 standard; peptide; 7 AA.	
XX		
XX	AAM44584;	
XX		
XX	25-OCT-2001 (first entry)	
XX		
XX	H11 binding site consensus conforming peptide (CCP) #855.	
XX		

KM Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KM immunogenically cross-reactive; cancer; immunogenic cancer cell;
KM cytostatic; vaccine; tumour-specific immunogenic response inducer;
KM astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;
KM ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
PI
XX WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 104; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (II) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;
XX
Query Match 100.0%; Score 22; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Gaps 0;
Matches 4; Conservative 0; Indels 0;
QY 1 KKFG 4
Db 2 KKFG 5
XX
RESULT 9
AAM45903
ID AAM45903 standard; peptide; 7 AA.
XX
AC AAM45903;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #2174.
XX
KM Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KM immunogenically cross-reactive; cancer; immunogenic cancer cell;
KM cytostatic; vaccine; tumour-specific immunogenic response inducer;
KM astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;
KM ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX
OS Homo sapiens.
OS Synthetic.
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
PI
XX WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 108; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (II) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;
XX
Query Match 100.0%; Score 22; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Gaps 0;
Matches 4; Conservative 0; Indels 0;
QY 1 KKFG 4
Db 2 KKFG 5
XX
RESULT 10
AAM4641
ID AAM4641 standard; peptide; 7 AA.
XX
AC AAM4641;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #912.
XX
KM Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KM immunogenically cross-reactive; cancer; immunogenic cancer cell;
KM cytostatic; vaccine; tumour-specific immunogenic response inducer;
KM astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;
KM ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
PN CA2290722-A1.

XX 08-JUN-2001.
PD 08-DEC-1999; 99CA-02290722.
XX 08-DEC-1999; 99CA-02290722.
XX 08-DEC-1999; 99CA-02290722.
XX (NOVO-) NOVOPHARM BIOTECH INC.
XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, MacDonald GC;
XX WPI; 2001-425937/46.
XX
XX Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
XX Example 4; Page 104; 154pp; English.
XX
XX The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumors that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPCCs specific to target cancer (TC). Also described is an isolated
CC or a population of different SPCCs consisting of immunogenic cancer cell
CC surface-associated SPCC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumor-specific immunogenic response
CC inducer. (I) is useful for treating 71 types of cancers or tumors in a
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumor (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 22; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKFg 4
Db 2 KKFg 5
RESULT 11
AAB35204
ID AAB35204 standard; peptide; 8 AA.
XX
XX AAB35204;
XX
XX 24-APR-2001 (first entry)
XX
XX Retroviral recombination assay peptide fragment #2.
XX
XX Retroviral recombination assay; gene therapy vector; viral vector; gag;
XX pol; replication; HIV.
XX
XX Unidentified.
XX
XX WO200104360-A2.
XX
XX 18-JAN-2001.
XX
XX 06-JUL-2000; 2000WO-US018597.
XX
XX 09-JUL-1999; 99US-0143015P.
XX 10-NOV-1999; 99US-0164626P.
XX
XX (UABR-) UAB RES FOUND.

PA (TRAN-) TRANZYME INC.
XX Kappes JC, Wu X, Wakefield J;
XX
XX WPI; 2001-091927/10.
XX N-PSDB; AAF24357.
XX
XX Retroviral recombination assays, systems and cells, useful for evaluating
PT the risk of producing a replication-competent retrovirus from a
PT retroviral-based vector.
XX
XX Disclosure; Fig 2; 88pp; English.
XX
XX The present invention describes a method for detecting a retroviral
CC genetic recombinant having gag and pol functions, involving a cell
CC suspected of having a recombinant and propagating the recombinant in the
CC presence of any necessary helper functions. This is useful as it enables
CC the detection of vectors which are able to replicate in the host cell,
CC and allows the production of vectors suitable for gene therapy. The
CC present sequence is a peptide fragment used to demonstrate the method of
CC the invention
XX
SQ Sequence 8 AA;
Query Match 100.0%; Score 22; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKFg 4
Db 5 KKFg 8
RESULT 12
AAG84320
ID AAG84320 standard; peptide; 10 AA.
XX
XX AAG84320;
XX
XX 11-SEP-2001 (first entry)
XX
XX Arabidopsis thaliana peptide ligand #960.
XX
XX Plant; peptide pesticide; peptide herbicide; agricultural research.
XX
XX Arabidopsis thaliana.
XX
XX WO200142279-A2.
XX
XX 14-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-GB004781.
XX
XX 13-DEC-1999; 99GB-00029469.
XX
XX (PROT-) PROTEOM LTD.
XX
XX Roberts GM, Heal JR;
XX
XX WPI; 2001-381629/40.
XX
XX
XX A set of peptide ligands for agricultural research and development of
PT therapeutic agents comprise specific complementary peptides to proteins
PT encoded by genes of plant genomes.
XX
XX Example 4; Page 165; 201pp; English.
XX
XX The present invention relates to a set of peptide ligands consisting of
CC specific complementary peptides to proteins encoded by genes of plant
CC genomes. The present sequence is one such peptide from Arabidopsis
CC thaliana. The peptides of the present invention are useful in an assay to
CC identify a peptide, especially a peptide pesticide or herbicide. The
CC peptides are also useful for tools for agricultural research and

CC development
XX
SQ Sequence 10 AA;

Query Match
Best Local Similarity 100.0%; Score 22; DB 4; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 1 KKF 4

RESULT 13
AAG6341
ID AAG6341 standard; peptide; 10 AA.

AC AAG6341;
XX 18-SEP-2001 (first entry)

DE Human complementary peptide, SEQ ID NO: 2535.

KW Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

PN WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB004776.

PR 13-DEC-1999; 99GB-00029464.

XX (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
XX candidates or pro-drugs.

PS Example 4; Page 404; 646pp; English.

XX The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification

XX Sequence 10 AA;

Query Match
Best Local Similarity 100.0%; Score 22; DB 4; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 3 KKF 6

RESULT 14
AAG6245
ID AAG6245 standard; peptide; 10 AA.

AC AAG6245;
XX

DT 18-SEP-2001 (first entry)

XX Human complementary peptide, SEQ ID NO: 2439.

KW Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

PN WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB004776.

PR 13-DEC-1999; 99GB-00029464.

XX (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
XX candidates or pro-drugs.

PS Example 4; Page 391; 646pp; English.

XX The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification

XX Sequence 10 AA;

Query Match
Best Local Similarity 100.0%; Score 22; DB 4; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 5 KKF 8

RESULT 15
AAG6183
ID AAG6183 standard; peptide; 10 AA.

AC AAG6183;
XX

DT 11-SEP-2001 (first entry)

DE Saccharomyces cerevisiae peptide, SEQ ID NO: 1132.

KW Saccharomyces cerevisiae; complementary peptide; peptide identification;
KM drug discovery; drug design.

OS Saccharomyces cerevisiae.

PN WO200142276-A1.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB004773.

PR 13-DEC-1999; 99GB-00029471.

XX (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-367863/38.

XX Identifying complementary peptides by analysis of protein and nucleotide
PT sequence databases, useful in drug design.

XX Example 3; Page 186; 488pp; English.

XX The invention relates to the identification of complementary peptides by
CC analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents and
CC drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC *Saccharomyces cerevisiae*

XX Sequence 10 AA;

Query Match 100.0%; Score 22; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
|||
Db 2 KKFG 5

Search completed: March 18, 2004, 11:16:22
Job time : 4.65427 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:13:05 ; Search time 0.463895 Seconds
(without alignments)
445.152 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KKFQ 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A COMB .pep:*
2: /cgn2_6/prodata/2/1aa/5B COMB .pep:*
3: /cgn2_6/prodata/2/1aa/6A COMB .pep:*
4: /cgn2_6/prodata/2/1aa/6B COMB .pep:*
5: /cgn2_6/prodata/2/1aa/6C COMB .pep:*
6: /cgn2_6/prodata/2/1aa/backfile1 .pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	12	1	US-08-260-582-15
2	22	100.0	12	5	PCT-US95-05471-15
3	22	100.0	13	1	US-08-465-325-135
4	22	100.0	13	4	US-08-115-737-135
5	22	100.0	15	2	US-08-338-882-5
6	22	100.0	16	4	US-08-914-999-20
7	22	100.0	17	1	US-08-132-767-11
8	22	100.0	18	1	US-08-233-203-10
9	22	100.0	18	2	US-08-338-882-18
10	22	100.0	18	2	US-08-338-882-20
11	22	100.0	19	1	US-08-132-767-8
12	22	100.0	19	1	US-08-132-767-10
13	22	100.0	20	1	US-07-711-183D-6
14	22	100.0	20	1	US-07-908-455A-74
15	22	100.0	20	1	US-08-132-767-7
16	22	100.0	20	1	US-08-434-120-11
17	22	100.0	20	1	US-08-465-325-11
18	22	100.0	20	3	US-08-782-997A-30
19	22	100.0	20	4	US-09-115-737-11
20	22	100.0	20	6	5254535-11
21	22	100.0	21	1	US-07-711-183D-5
22	22	100.0	21	1	US-07-965-663A-1
23	22	100.0	21	1	US-07-965-663A-2
24	22	100.0	21	1	US-07-965-663A-3
25	22	100.0	21	1	US-07-965-663A-22
26	22	100.0	21	1	US-07-908-455A-73
27	22	100.0	21	1	US-08-132-767-6

28	22	100.0	21	1	US-08-132-767-9	Sequence 9, Appl1
29	22	100.0	21	1	US-08-434-120-10	Sequence 10, Appl1
30	22	100.0	21	1	US-08-465-325-10	Sequence 10, Appl1
31	22	100.0	21	2	US-08-448-600-9	Sequence 9, Appl1
32	22	100.0	21	3	US-08-937-610-17	Sequence 17, Appl1
33	22	100.0	21	4	US-09-115-737-10	Sequence 10, Appl1
34	22	100.0	21	6	5254535-10	Patent No. 5254535
35	22	100.0	22	1	US-07-711-183D-3	Sequence 3, Appl1
36	22	100.0	22	1	US-07-965-663A-6	Sequence 6, Appl1
37	22	100.0	22	1	US-07-965-663A-19	Sequence 7, Appl1
38	22	100.0	22	1	US-07-965-663A-19	Sequence 19, Appl1
39	22	100.0	22	1	US-07-965-663A-20	Sequence 20, Appl1
40	22	100.0	22	1	US-07-965-663A-21	Sequence 21, Appl1
41	22	100.0	22	1	US-07-908-455A-71	Sequence 71, Appl1
42	22	100.0	22	1	US-07-908-455A-72	Sequence 72, Appl1
43	22	100.0	22	1	US-08-132-767-5	Sequence 5, Appl1
44	22	100.0	22	1	US-08-132-767-12	Sequence 12, Appl1
45	22	100.0	22	1	US-08-132-767-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-08-260-582-15
; Sequence 15, Application US/08260582
; Patent No. 563182
; GENERAL INFORMATION:
; APPLICANT: McCoy, John M.
; APPLICANT: Lu, Zhifan
; TITLE OF INVENTION: METHOD OF DETECTING LIGAND
; TITLE OF INVENTION: INTERACTIONS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/260,582
; FILING DATE: 16-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meinerdt, M. C.
; REGISTRATION NUMBER: 31,544
; REFERENCE/DOCKET NUMBER: GI 5236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-260-582-15

Query Match 100.0%; Score 22; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKFQ 4
||||

Db 6 KKFG 9

RESULT 2
PCT-US95-05471-15
Sequence 15, Application PC/TUS9505471
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHOD OF DETECTING LIGAND INTERACTIONS
NUMBER OF SEQUENCES: 76
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION FOR SEQ ID NO: 15:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-05471-15

Query Match 100.0%; Score 22; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
Db 6 KKFG 9

RESULT 3
US-08-465-325-135
Sequence 135, Application US/08465325
Patent No. 5686563
GENERAL INFORMATION:
APPLICANT: Magalini Pharmaceuticals Inc.
APPLICANT: 5110 Campus Drive
APPLICANT: Plymouth Meeting, PA 19462
TITLE OF INVENTION: Biologically Active Peptides Having
TITLE OF INVENTION: N-Terminal Substitutions
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I. Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,325
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B

REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0021-03000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 135:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-465-325-135

Query Match 100.0%; Score 22; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
Db 1 KKFG 4

RESULT 4
US-09-115-737-135
Sequence 135, Application US/09115737
Patent No. 6348445
GENERAL INFORMATION:
APPLICANT: U. Prasad Kari
Michael McLane
Taffy J. Williams
TITLE OF INVENTION: Biologically Active Peptides With Reduced
Toxicity In Animals and a Method for Preparing Same
NUMBER OF SEQUENCES: 156
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,737
FILING DATE: 15-Jul-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,330
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0021-06000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 135:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 135:

US-09-115-737-135

Query Match 100.0%; Score 22; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
11111
Db 1 KKEG 4

RESULT 5

US-08-338-882-5
Sequence 5, Application US/0833882
Patent No. 5912231

GENERAL INFORMATION:

APPLICANT: Houghton, Richard A.

APPLICANT: Cuervo, Julio H.

TITLE OF INVENTION: Substitution Analogues of Magainin

TITLE OF INVENTION: Peptides

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Goldsmith, Shore, Suter &

ADDRESSEE: Milamow, Ltd.

STREET: 180 No. 5912231th Stetson, Suite 4700

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/338.882

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/615,125

FILING DATE: 15-NOV-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gansson, Edward P

REGISTRATION NUMBER: 29,381

REFERENCE/DOCKET NUMBER: SCRF 186.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5400

TELEFAX: (312) 616-5460

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 15

OTHER INFORMATION: /note= "Xaa=alanine amide"

US-08-338-882-5

Query Match 100.0%; Score 22; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
11111
Db 10 KKEG 13

RESULT 6

US-08-914-999-20
Sequence 20, Application US/08914999
Patent No. 6346406

GENERAL INFORMATION:

APPLICANT: Ryazanov, Alexey G.

APPLICANT: Halt, William N.

APPLICANT: Pavur, Karen S.

TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)

TITLE OF INVENTION: AND METHODS OF USE THEREFOR

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.

STREET: 411 Hackensack Ave, Continental Plaza, 4th

STREET: Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/914,999

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 601-1-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-345-1684

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-914-999-20

Query Match 100.0%; Score 22; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
11111
Db 2 KKEG 5

RESULT 7

US-08-132-767-11
Sequence 11, Application US/08132767
Patent No. 5518912

GENERAL INFORMATION:

APPLICANT: Michael A. Zaslloff; Nicole Resnick

TITLE OF INVENTION: NOVEL ENDOPEPTIDASE

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912ris

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: WORDPERFECT 5.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/132,767

FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/002,109
FILING DATE:
APPLICATION NUMBER: US/07/685,723
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patricia A. Schreck
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: CH-0006
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 17
TYPE: Amino acids
TOPOLOGY: Unknown
US-08-132-767-11

Query Match 100.0%; Score 22; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
Db 10 KKFg 13

RESULT 8
US-08-233-203-10

Sequence 10, Application US/08233203
Patent No. 5409898

GENERAL INFORMATION:

APPLICANT: Darveau, Richard P.

APPLICANT: Blake, James J.

APPLICANT: Cosand, Wesley L.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INJECTIONS CAUSED BY ORGANISM SENSITIVE TO BETA-LACTAM

TITLE OF INVENTION: ANTIBIOTICS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bristol-Myers Squibb Company, Patent

STREET: 3005 First Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/233,203

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/655,321

FILING DATE: 19-FEB-1991

APPLICATION NUMBER: US 07/484,020

FILING DATE: 23-FEB-1990

ATTORNEY/AGENT INFORMATION:

NAME: Poor, Brian W.

REGISTRATION NUMBER: 32,928

REFERENCE/DOCKET NUMBER: ON0063A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206/448-4775

TELEFAX: 206/448-4800

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
US-08-233-203-10

Query Match 100.0%; Score 22; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
Db 15 KKFg 18

RESULT 9

US-08-338-882-18

Sequence 18, Application US/08338882

Patent No. 5912231

GENERAL INFORMATION:

APPLICANT: Houghten, Richard A.

APPLICANT: Querio, Julio H.

TITLE OF INVENTION: Substitution Analogues of Magainin

TITLE OF INVENTION: Peptides

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &

STREET: 180 No. 5912231th Stetson, Suite 4700

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/338,882

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/615,125

FILING DATE: 15-NOV-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gamsen, Edward P.

REGISTRATION NUMBER: 29,381

REFERENCE/DOCKET NUMBER: SCRF 186.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5400

TELEFAX: (312) 616-5460

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-338-882-18

Query Match 100.0%; Score 22; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
Db 6 KKFg 9

RESULT 10

US-08-338-882-20

Sequence 20, Application US/08338882

Patent No. 5912231

GENERAL INFORMATION:
APPLICANT: Houghten, Richard A.
TITLE OF INVENTION: Substitution Analogues of Magastatin
TITLE OF INVENTION: Peptides
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Shore, Suter &
ADDRESSEE: Milnamow, Ltd.
STREET: 180 No. 591231th Stetson, Suite 4700
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,882
FILING DATE:
CLASSIFICATION: 514
APPLICATION NUMBER: US 07/615,125
FILING DATE: 15-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gansson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: SCRF 186.1
TELEPHONE: (312) 616-5400
TELEFAX: (312) 616-5460
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-338-882-20

Query Match 100.0%; Score 22; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 6 KKF 9

RESULT 11
US-08-132-767-8
Sequence 8, Application US/08132767
Patent No. 5518912
GENERAL INFORMATION:
APPLICANT: Michael A. Zasloff; Nicole Resnick
TITLE OF INVENTION: NOVEL ENDOPEPTIDASE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/132,767
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/002,109
FILING DATE:
APPLICATION NUMBER: US/07/685,723
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patricia A. Schreck
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: CH-0006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 19
TYPE: Amino acids
TOPOLOGY: Unknown
US-08-132-767-8

Query Match 100.0%; Score 22; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 6 KKF 9

RESULT 12
US-08-132-767-10
Sequence 10, Application US/08132767
Patent No. 5518912
GENERAL INFORMATION:
APPLICANT: Michael A. Zasloff; Nicole Resnick
TITLE OF INVENTION: NOVEL ENDOPEPTIDASE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/132,767
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/002,109
FILING DATE:
APPLICATION NUMBER: US/07/685,723
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patricia A. Schreck
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: CH-0006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19
TYPE: Amino acids
TOPOLOGY: Unknown
US-08-132-767-10

Query Match 100.0%; Score 22; DB 1; Length 19;

Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
|||
Db 10 KKFG 13

RESULT 13
US-07-711-183D-6
; Sequence 6, Application US/07711183D
; Patent No. 5254537
; GENERAL INFORMATION:
; APPLICANT: Zaslloff, Michael
; TITLE OF INVENTION: Composition and Treatment with
; TITLE OF INVENTION: Peptide Combinations
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/711,183D
; FILING DATE: 19910510
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07 346894
; FILING DATE: 03-MAY-1989
; APPLICATION NUMBER: US 07 302985
; FILING DATE: 30-JAN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 421250-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 6
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: magainin peptide.
; PUBLICATION INFORMATION:
; AUTHORS: Zaslloff, Michael
; JOURNAL: Proceedings of the National Academy
; OF SCIENCES
; VOLUME: 84
; PAGES: 5449-5453
; DATE: AUG - 1987
; DOCUMENT NUMBER: US 4810777
; FILING DATE: 04-MAR-1987
; PUBLICATION DATE: 07-MAR-1989
; US-07-711-183D-6

Query Match 100.0%; Score 22; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
|||
Db 7 KKFG 10

RESULT 14
US-07-908-455A-74
; Sequence 74, Application US/07908455A
; Patent No. 5459237
; GENERAL INFORMATION:
; APPLICANT: Berkowitz, Barry A.
; APPLICANT: Kari, U. Prasad
; APPLICANT: Maloy, W. Lee
; TITLE OF INVENTION: No. 5459237el Peptide Compositions and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,455A
; FILING DATE: 19920702
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07686115
; FILING DATE: 15-APR-1991
; APPLICATION NUMBER: US 07476629
; FILING DATE: 08-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 421250-122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: magainin peptide.
; OTHER INFORMATION: amide- or carboxy- terminated
; PUBLICATION INFORMATION:
; AUTHORS: Zaslloff, Michael
; JOURNAL: Proc. Nat. Acad. Sci.
; VOLUME: 84
; PAGES: 5449-5453
; DATE: AUG - 1987
; DOCUMENT NUMBER: US 4810777
; FILING DATE: 04-MAR-1987
; PUBLICATION DATE: 07-MAR-1989
; US-07-908-455A-74

Query Match 100.0%; Score 22; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
|||
Db 7 KKFG 10

RESULT 15

US-08-132-767-7
; Sequence 7, Application US/08132767
; Patent No. 5518912
; GENERAL INFORMATION:
; APPLICANT: Michael A. Zaeloff; Nicole Resnick
; TITLE OF INVENTION: NOVEL ENDOPEPTIDASE
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912z1s
; STREET: One Liberty Place - 46th Floor
; City: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/132,767
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/002,109
; FILING DATE:
; APPLICATION NUMBER: US/07/685,723
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Patricia A. Schreck
; REGISTRATION NUMBER: 33,777
; REFERENCE/DOCKET NUMBER: CH-0006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: Amino acids
; TOPOLOGY: Unknown
; US-08-132-767-7

Query Match 100.0%; Score 22; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKG 4
|||
Db 7 KKG 10

Search completed: March 18, 2004, 11:21:22
Job time: 1.46389 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:19:25 ; Search time 1.19037 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-3
Perfect score: 22
Sequence: 1 KXFG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues
Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	4	13	US-10-076-421-3
2	22	100.0	10	10	US-09-572-404B-2439
3	22	100.0	10	10	US-09-572-404B-2535
4	22	100.0	10	10	US-09-572-270A-960
5	22	100.0	10	10	US-09-573-823C-328
6	22	100.0	16	9	US-09-994-485-20
7	22	100.0	22	9	US-09-807-120-3
8	22	100.0	22	10	US-09-820-053A-24
9	22	100.0	22	10	US-09-904-753-3
10	22	100.0	22	10	US-09-904-753-4
11	22	100.0	22	14	US-10-109-171-24
12	22	100.0	23	9	US-09-030-619-211
13	22	100.0	23	10	US-09-820-053A-7
14	22	100.0	23	10	US-09-904-753-2
15	22	100.0	23	10	US-09-798-026B-14

16	22	100.0	23	14	US-10-109-171-7	Sequence 7, Appli
17	22	100.0	23	15	US-10-252-773-13	Sequence 13, Appli
18	22	100.0	23	15	US-10-277-233-211	Sequence 211, App
19	22	100.0	38	12	US-10-424-599-165729	Sequence 165729,
20	22	100.0	41	12	US-10-424-599-270907	Sequence 270907,
21	22	100.0	43	14	US-10-097-111-475	Sequence 475, App
22	22	100.0	43	14	US-10-321-857-65	Sequence 65, Appli
23	22	100.0	43	14	US-10-318-675-65	Sequence 1, Appli
24	22	100.0	44	12	US-10-349-543-1	Sequence 144596,
25	22	100.0	46	12	US-10-424-599-144596	Sequence 144596,
26	22	100.0	46	12	US-10-424-599-173819	Sequence 223867,
27	22	100.0	46	12	US-10-424-599-223867	Sequence 223867,
28	22	100.0	46	12	US-10-424-599-236167	Sequence 236167,
29	22	100.0	47	9	US-09-880-503-2	Sequence 2, Appli
30	22	100.0	48	12	US-10-424-599-250270	Sequence 250270,
31	22	100.0	49	9	US-09-864-761-33472	Sequence 33472, A
32	22	100.0	49	12	US-10-424-599-208329	Sequence 208329,
33	22	100.0	49	14	US-10-349-543-5	Sequence 5, Appli
34	22	100.0	50	9	US-09-864-761-35403	Sequence 35403, A
35	22	100.0	50	12	US-10-424-599-272294	Sequence 272294,
36	22	100.0	51	9	US-09-925-301-1679	Sequence 1679, Ap
37	22	100.0	51	9	US-09-250-611-43	Sequence 43, Appli
38	22	100.0	51	12	US-10-424-599-249968	Sequence 249968,
39	22	100.0	51	12	US-10-424-599-283379	Sequence 283379,
40	22	100.0	52	12	US-10-424-599-170813	Sequence 170813,
41	22	100.0	53	11	US-09-833-245-996	Sequence 996, App
42	22	100.0	54	9	US-09-864-761-37535	Sequence 37535, A
43	22	100.0	55	12	US-10-424-599-235683	Sequence 235683,
44	22	100.0	57	14	US-10-029-386-31376	Sequence 31376, A
45	22	100.0	57	12	US-10-424-599-209467	Sequence 209467,

ALIGNMENTS

RESULT 1
US-10-076-421-3
; Sequence 3, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-3

Query Match 100.0%; Score 22; DB 13; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KXFG 4
Db 1 KXFG 4
RESULT 2
US-09-572-404B-2439
; Sequence 2439, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd

```

; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2439
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 396-405 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-09-572-404B-2439

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 5 KKF 8

RESULT 3
US-09-572-404B-2535
; Sequence 2535, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2535
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 398-407 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-09-572-404B-2535

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 3 KKF 6

RESULT 4
US-09-572-270A-960
; Sequence 960, Application US/09572270A
; Publication No. US20030148368A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Inter- complementary peptide listing
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/572,270A
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 1144
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 960
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
; OTHER INFORMATION: Sequence located in RPS16. at 130-139 and may interact with
US-09-572-270A-960
```

```

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 1 KKF 4

RESULT 5
US-09-573-822C-328
; Sequence 328, Application US/09573822C
; Publication No. US20030199011A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands generated from microbial genome seq
; FILE REFERENCE: Microbe patent
; CURRENT APPLICATION NUMBER: US/09/573,822C
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 804
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 328
; LENGTH: 10
; TYPE: PRT
; ORGANISM: mycoplasma genitalium
; FEATURE:
; OTHER INFORMATION: Sequence located in MG214 at 76-85 and may interact with Sequence
; OTHER INFORMATION: this patent.
US-09-573-822C-328

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 1 KKF 4

RESULT 6
US-09-994-485-20
; Sequence 20, Application US/09994485
; Patent No. US20020142429A1
; GENERAL INFORMATION:
; APPLICANT: Ryazanov, Alexey G.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESS: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/994,485
; FILING DATE: 27-Nov. US20020142429A1-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/914,999
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq, David A.
; REGISTRATION NUMBER: 26,742
```

REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-994-485-20

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 16;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 2 KKF 5

RESULT 7
US-09-807-720-3
Sequence 3, Application US/09807720
Patent No. US20020162135A1
GENERAL INFORMATION:
APPLICANT: DANIEL, HENRY
TITLE OF INVENTION: EXPRESSION OF AN ANTIMICROBIAL PEPTIDE VIA THE PLASTID
FILE REFERENCE: 1462-PCT-US-00
CURRENT APPLICATION NUMBER: US/09/807,720
CURRENT FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/185,662
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-807-720-3

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 10 KKF 13

RESULT 8
US-09-820-053A-24
Sequence 24, Application US/09820053A
Publication No. US20030083243A1
GENERAL INFORMATION:
APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
FILE REFERENCE: HELIX27
CURRENT APPLICATION NUMBER: US/09/820,053A
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 165
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 24
LENGTH: 22
TYPE: PRT
ORGANISM: ARTIFICIAL SEQUENCE

FEATURE:
OTHER INFORMATION: SYNTHETIC SEQUENCE
NAME/KEY: MOD RES
LOCATION: (22)
OTHER INFORMATION: AMIDATION
US-09-820-053A-24

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 10 KKF 13

RESULT 9
US-09-904-753-3
Sequence 3, Application US/09904753
Publication No. US20030092612A1
GENERAL INFORMATION:
APPLICANT: Lynos, Robert T
TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
Emulsions, and Suspensions
FILE REFERENCE: 2973 ver 2
CURRENT APPLICATION NUMBER: US/09/904,753
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: WO 96/25183
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (22)
OTHER INFORMATION: Xaa at position 22 is Lys-amide
US-09-904-753-3

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 10 KKF 13

RESULT 10
US-09-904-753-4
Sequence 4, Application US/09904753
Publication No. US20030092612A1
GENERAL INFORMATION:
APPLICANT: Lynos, Robert T
TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
Emulsions, and Suspensions
FILE REFERENCE: 2973 ver 2
CURRENT APPLICATION NUMBER: US/09/904,753
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: WO 96/25183
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: maginin analog
US-09-904-753-4

Query Match 100.0%; Score 22; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 11
US-10-109-171-24
; Sequence 24, Application US/10109171
; Publication No. US20030109452A1
; GENERAL INFORMATION:
; APPLICANT: Owen, Donald R.
; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES AND METHODS FOR THEIR USE
; FILE REFERENCE: HELIX28
; CURRENT APPLICATION NUMBER: US/10/109,171
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 22
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC SEQUENCE
; NAME/KEY: MOD RES
; LOCATION: (22)
; OTHER INFORMATION: AMIDATION
US-10-109-171-24

Query Match 100.0%; Score 22; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 12
US-09-030-619-211
; Sequence 211, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erile, Douglas
; APPLICANT: Frazer, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 211
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-030-619-211

Query Match 100.0%; Score 22; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 13
US-09-820-053A-7
; Sequence 7, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:
; APPLICANT: Owen, Donald R.
; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
; FILE REFERENCE: HELIX27
; CURRENT APPLICATION NUMBER: US/09/820,053A
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 23
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC SEQUENCE
; NAME/KEY: MOD RES
; LOCATION: (23)
; OTHER INFORMATION: AMIDATION
US-09-820-053A-7

Query Match 100.0%; Score 22; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 14
US-09-904-753-2
; Sequence 2, Application US/09904753
; Publication No. US20030092612A1
; GENERAL INFORMATION:
; APPLICANT: Lynos, Robert T
; TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
; TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
; TITLE OF INVENTION: Emulsions, and Suspensions
; FILE REFERENCE: 2973 ver 2
; CURRENT APPLICATION NUMBER: US/09/904,753
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: WO 96/25183
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-904-753-2

Query Match 100.0%; Score 22; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 15
US-09-798-026B-14
; Sequence 14, Application US/09798026B
; Publication No. US20030148936A1
; GENERAL INFORMATION:

```

; APPLICANT: ALPHARMA AS
; TITLE OF INVENTION: Cytotoxic Peptides Modified by Bulky or Lipophilic Moieties
; FILE REFERENCE: 40745-2
; CURRENT APPLICATION NUMBER: US/09/798,026B
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: GB 9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 23
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(23)
; OTHER INFORMATION: Artificial peptide = native magainin peptide or mutated native
; OTHER INFORMATION: magainin peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)..(6)
; OTHER INFORMATION: position of possible substitution (L-->W)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)..(8)
; OTHER INFORMATION: position of possible substitution (S-->W)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (19)..(19)
; OTHER INFORMATION: position of possible substitution (E-->W)
; US-09-798-026B-14

```

```

Query Match          100.0%; Score 22; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 KKFG 4
        ||||
Db      10 KKFG 13

```

Search completed: March 18, 2004, 11:36:07
 Job time: 1.44037 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 0.41179 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KKEG 4

Scoring table: BLOSUM62
Gap 10.0, Gape 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	37	1 HSWT93	histone H2A.3 - wh
2	22	100.0	43	2 A39313	chaperonin, 58K -
3	22	100.0	50	2 E95014	hypothetical prote
4	22	100.0	50	2 G97887	orf151 (imported) -
5	22	100.0	58	2 H83721	hypothetical prote
6	22	100.0	61	1 R3YM14	ribosomal protein
7	22	100.0	61	2 D98277	hypothetical prote
8	22	100.0	62	2 A65045	hypothetical prote
9	22	100.0	67	2 AE3277	LEU ribosomal prot
10	22	100.0	83	2 S46930	leg292 protein - m
11	22	100.0	84	2 G81425	50S ribosomal prot
12	22	100.0	84	2 E90005	ribosomal protein
13	22	100.0	86	2 F97061	hypothetical prote
14	22	100.0	87	2 P69542	conserved hypotet
15	22	100.0	88	2 A64557	ribosomal protein
16	22	100.0	88	2 E71951	ribosomal protein
17	22	100.0	88	2 T44265	ribosomal protein
18	22	100.0	88	2 T43610	probable IS1617 tr
19	22	100.0	89	2 AG2918	50S ribosomal prot
20	22	100.0	89	2 B97693	ribosomal protein
21	22	100.0	89	2 E87288	ribosomal protein
22	22	100.0	89	2 T03669	reverse transcript
23	22	100.0	90	2 T03671	reverse transcript
24	22	100.0	91	2 E75560	ribosomal protein
25	22	100.0	91	2 T23795	hypothetical prote
26	22	100.0	92	2 T03615	retrovirus-related
27	22	100.0	92	2 T03666	reverse transcript
28	22	100.0	94	2 S77047	transposase srr17
29	22	100.0	99	2 T32862	hypothetical prote

30	22	100.0	100	2 T11953	ribosomal protein
31	22	100.0	100	2 E69460	conserved hypotet
32	22	100.0	107	2 JQ1742	hypothetical 12.6K
33	22	100.0	108	2 AG3005	conserved hypotet
34	22	100.0	111	2 T13343	hypothetical prote
35	22	100.0	115	2 T40621	hypothetical prote
36	22	100.0	117	2 G64335	hypothetical prote
37	22	100.0	117	2 H84651	hypothetical prote
38	22	100.0	118	1 PSKFR3	phospholipase A2 (
39	22	100.0	118	1 PSKFR2	phospholipase A2 (
40	22	100.0	118	1 PSKFR1	phospholipase A2 (
41	22	100.0	119	2 S74925	transposase s11065
42	22	100.0	119	2 S75488	transposase s1x211
43	22	100.0	119	2 S74836	transposase s1x085
44	22	100.0	119	2 S75827	transposase s1x128
45	22	100.0	119	2 S75590	transposase s11125

ALIGNMENTS

RESULT 1

HSWT93
histone H2A.3 - wheat (fragment)
C/Species: Triticum aestivum (common wheat)
C/Date: 31-May-1980 #sequence_revision 31-May-1980 #text_change 16-Feb-1997
C/Accession: A02604
R/Rodriguez, J.A.; Brandt, W.F.; von Holt, C.
Biochim. Biophys. Acta 578, 196-206, 1979
A/Title: Plant histone 2 from wheat germ, a family of histone H2A variants.
A/Reference number: A90629; MUID:79209696; PMID:454665
A/Accession: A02604
A/Molecule type: protein
A/Residues: 1-37 <ROD>
A/Experimental source: germ
C/Suprafamily: histone H2A
C/Keywords: acetylated amino end; chromosomal protein; DNA binding; nucleosome core
F/1/Modified site: acetylated amino end (Met) #status experimental

Query Match 100.0%; Score 22; DB 1; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4

DB 12 KKEG 15

RESULT 2

A39313
chaperonin, 58K - Thermus aquaticus (fragment)
C/Species: Thermus aquaticus
C/Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 12-Sep-1997
C/Accession: A39313
R/Raguchi, H.; Konishi, J.; Ishii, N.; Yoshida, M.
J. Biol. Chem. 266, 22411-22418, 1991
A/Title: A chaperonin from a thermophilic bacterium, Thermus thermophilus, that controls
A/Reference number: A39313; MUID:92042183; PMID:1662319
A/Accession: A39313
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-43 <TAG>
C/Suprafamily: chaperonin groEL
C/Keywords: molecular chaperone

Query Match 100.0%; Score 22; DB 2; Length 43;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4

DB 40 KKEG 43

RESULT 3
E95014
hypothetical protein SP0124 [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C/Accession: E95014
R/Tetcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
neon, T.; Hickey, E.R.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Lofthus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A./Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357203; PMID:115463916
A/Accession: E95014
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-50 <KUR>
A/Cross-references: GB:AE005672; PIDN:AAK74310.1; PID:g14971592; GSPDB:GN00164; TIGR:SP4
A/Experimental source: strain TIGR4
A/Gene: SP0124

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 44 KKF 47

RESULT 4
G97887
orf51 [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C/Accession: G97887
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
y, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, F.; McAhlem, S.; M
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: G97887
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-50 <KUR>
A/Cross-references: GB:AE007317; PIDN:AAK98931.1; PID:g15457666; GSPDB:GN00174
A/Gene: orf51

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 44 KKF 47

RESULT 5
H83721
hypothetical protein BH0576 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: H83721
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: H83721

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-58 <STO>
A/Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA04295.1; GSPDB:GN001
A/Experimental source: strain C-125
A/Gene: BH0576

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 30 KKF 33

RESULT 6
R3YM14
ribosomal protein S14 - Mycoplasma capricolum
C/Species: Mycoplasma capricolum
C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 07-Dec-1999
C/Accession: S02844
R/Ohkubo, S.; Muto, A.; Kawachi, Y.; Yamao, F.; Oosawa, S.
Mol. Gen. Genet. 210, 314-322, 1987
A/Title: The ribosomal protein gene cluster of Mycoplasma capricolum.
A/Reference number: S02830; MUID:88142549; PMID:3481422
A/Accession: S02844
A/Molecule type: DNA
A/Residues: 1-61 <OHK>
A/Cross-references: EMBL:X06414; NID:g44207; PIDN:CAA29717.1; PID:g44222
A/Gene: rps14
A/Gene code: SGC3
C/Superfamily: Escherichia coli ribosomal protein S14
C/Keywords: protein biosynthesis; ribosome

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 35 KKF 38

RESULT 7
D98277
hypothetical protein AGR_L_2351 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C/Accession: D98277
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurolto, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A87359; MUID:21608551; PMID:11743194
A/Accession: D98277
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-61 <KUR>
A/Cross-references: GB:AE007870; PIDN:AAK89742.1; PID:g15159660; GSPDB:GN00170
A/Gene: AGR_L_2351
A/Map position: linear chromosome

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 32 KKF 35

RESULT 8
A:Accession: A65045
hypothetical protein b2656 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: A65045
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; PMID:97426617; PMID:9278503
A:Accession: A65045
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-62 <KUR>
A:Cross-references: GB:AE000350; GB:U00096; NID:G2367147; PIDN:AACT5703.1; PID:gl789009;
A:Experimental source: strain K-12, substrain M01655

Query Match 100.0%; Score 22; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
DB 15 KXFG 18

RESULT 9

AE3277
LSU ribosomal protein L27P [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 18-Nov-2002
C:Accession: AE3277
R:DelVecchio, V.G.; Kaputrl, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldman, E.; Selkov, E.; Blizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL51384.1; PID:gl7982087; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME10202
A:Map position: 1
C:Superfamily: Escherichia coli ribosomal protein L27, eubacterial ribosomal protein L27

Query Match 100.0%; Score 22; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
DB 2 KXFG 5

RESULT 10

S46930
teg292 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S46930
R: Lopez-Fernandez, L.A.; del Mazo, J.
submitted to the EMBL Data Library, March 1994
A:Reference number: S46929
A:Accession: S46930
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-83 <LOP>

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-83 <LOP>

A:Cross-references: EMBL:X80433; NID:g515352; PIDN:CAA56628.1; PID:g939954

Query Match 100.0%; Score 22; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
DB 43 KXFG 46

RESULT 11

G81425
50S ribosomal protein L27 Cj0095 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Nov-2002
C:Accession: G81425
R:Perkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett)
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; PMID:20150912; PMID:10688204
A:Accession: G81425
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84 <PAR>
A:Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB72579.1; PID:g696758;
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: rpmA; Cj0095
C:Superfamily: Escherichia coli ribosomal protein L27, eubacterial ribosomal protein L27

Query Match 100.0%; Score 22; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
DB 24 KXFG 27

RESULT 12

E90005
ribosomal protein L31 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002
C:Accession: E90005
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; i
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: E90005
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84 <KUR>
A:Cross-references: GB:BA000018; PID:g13701914; PIDN:BA43206.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: rpmE
C:Superfamily: Escherichia coli ribosomal protein L31

Query Match 100.0%; Score 22; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
DB 76 KXFG 79

RESULT 13

P97061
hypothetical protein CAC1312 [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C/Accession: F97061
R/Molling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A/Reference number: A96900; MUID:2135325; PMID:21359325
A/Accession: F97061
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-86 <KIR>
A/Cross-references: GB:AE001437; PIDN:AAK79281.1; PID:GL5024241; GSPDB:GN00168
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC1312

Query Match 100.0%; Score 22; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
|||
Db 75 KKF 78

RESULT 14
F69542
conserved hypothetical protein AF2342 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: F69542
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Uteerback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.F.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: F69542
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-87 <KLE>
A/Cross-references: GB:AE000943; GB:AE000782; NID:G2689266; PIDN:AA88914.1; PID:G264817
C/Superfamily: hypothetical protein MG0071

Query Match 100.0%; Score 22; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
|||
Db 23 KKF 26

RESULT 15
A64557
ribosomal protein L27 - Helicobacter pylori (strain 26695)
C/Species: Helicobacter pylori
C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 13-Aug-1999
C/Accession: A64557
R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalaf, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, L.; Nature 388, 539-547, 1997
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A/Reference number: A64520; MUID:97394467; PMID:9252185
A/Accession: A64557
A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
A/Residues: 1-88 <TOM>
A/Cross-references: GB:AE000548; GB:AE000511; NID:G2313391; PIDN:AA07366.1; PID:G2313391
C/Superfamily: Escherichia coli ribosomal protein L27; eubacterial ribosomal protein L27
F/2-84/Domain: eubacterial ribosomal protein L27 homology <L27>

Query Match 100.0%; Score 22; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
|||
Db 24 KKF 27

Search completed: March 18, 2004, 11:20:17
Job time: 1.41138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 10:54:45 ; Search time 0.245077 Seconds

(without alignments)
849.859 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KRFK 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	22	100.0	60	RS14_MYCCA	P10130 mycoplasma
2	22	100.0	61	RS14_CLOPE	Q8XHC6 clostridium
3	22	100.0	76	APM2_HUMAN	Q15847 homo sapien
4	22	100.0	84	R31B_PHOL	Q7NA98 photorhabd
5	22	100.0	84	R31B_STAM	Q9A688 campylobact
6	22	100.0	84	RL27_CAMJE	Q9P311 campylobact
7	22	100.0	85	R31B_STAMP	Q8CRA8 campylobact
8	22	100.0	88	RL27_HELPU	Q9ZM68 helicobacte
9	22	100.0	88	RL27_HELPU	P56050 helicobacte
10	22	100.0	88	RL27_SYNP7	Q9Z3H6 synechococc
11	22	100.0	89	RL27_AGRF5	Q8UBH6 agrobacteri
12	22	100.0	89	RL27_BRUAB	Q8VW58 bruceella ab
13	22	100.0	89	RL27_BRUME	Q8YJ84 bruceella me
14	22	100.0	89	RL27_BRUSU	Q8FY18 bruceella su
15	22	100.0	89	RL27_CAVCR	Q9ABH3 caulobacter
16	22	100.0	89	RL27_RHILC	Q9ABX0 rhizobium
17	22	100.0	89	RL27_RHILC	Q921B7 rhizobium
18	22	100.0	91	RL27_RHILC	Q21557 caenorhabdi
19	22	100.0	91	RL27_DEIRA	Q9RY65 deinococcus
20	22	100.0	96	Y12A_BP74	R39493 bacterioph
21	22	100.0	100	RK27_CVACA	R19855 cyanidium c
22	22	100.0	103	CHLB_ZAMFI	P37857 zamia fisch
23	22	100.0	103	DL4A_HUMAN	Q96015 homo sapien
24	22	100.0	105	DL4A_MOUSE	Q9DCM4 mus musculu
25	22	100.0	111	RNPA_FUSNN	Q8RHA6 fusobacteri
26	22	100.0	115	PT19_STYPL	P28211 styela plic
27	22	100.0	117	Y286_METUA	O57734 methanococc
28	22	100.0	118	PA26_BUNFA	P06627 bungarus fa
29	22	100.0	118	PA26_BUNFA	P06628 bungarus fa
30	22	100.0	118	PA26_BUNFA	P06629 bungarus fa
31	22	100.0	119	RNPA_CHIMU	Q9P177 chlamydia m
32	22	100.0	120	RNPA_CHLTR	Q84789 chlamydia t
33	22	100.0	122	RL7_BUCAP	P41188 buchnera ap

34	22	100.0	122	1	RL7_BUCAP	Q89B19 buchnera ap
35	22	100.0	123	1	RL7_WIGBR	Q8D234 wigglewort
36	22	100.0	125	1	ACPS_NEIMA	Q9TQW2 neisseria m
37	22	100.0	125	1	ACPS_NEIMA	Q9TQW2 neisseria m
38	22	100.0	125	1	RL7_HELPU	P55834 helicobacte
39	22	100.0	130	1	RS8_METUA	P54041 methanococc
40	22	100.0	131	1	RS9_MYCCA	Q7NB14 mycoplasma
41	22	100.0	132	1	RS9_MYCCA	P47656 mycoplasma
42	22	100.0	132	1	RS9_MYCCA	P75179 mycoplasma
43	22	100.0	133	1	RS9_UREPA	Q9PPC3 ureaplasma
44	22	100.0	134	1	RS9_METKA	Q8TVB5 methanopyru
45	22	100.0	135	1	RS16_PIG	Q29201 sus scrofa

ALIGNMENTS

RESULT 1
ID RS14_MYCCA STANDARD; PRT; 60 AA.
AC P10130;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE 30S ribosomal protein S14.
GN RPSN.
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2095;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27343 / K1d;
RX MEDLINE=88142549; PubMed=3481422;
RA Okubo S., Muto A., Kawachi Y., Yamao F., Ogasawa S.;
RT "The ribosomal protein gene cluster of Mycoplasma capricolum.";
RL Mol. Gen. Genet. 210:314-322(1987).
CC -1- FUNCTION: Known to be required for the assembly of 30S particles
CC and may also be responsible for determining the conformation of
CC the 16S rRNA at the A site (By similarity).
CC -1- SIMILARITY: Belongs to the S14p family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X06414; CAA29717.1; ALT_TERM.
CC PIR; S02844; R3YM14.
CC DR InterPro: IPR001209; Ribosomal_S14.
CC DR Pfam: PF00253; Ribosomal_S14; 1.
CC DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
CC KW Ribosomal protein.
CC
CC SEQUENCE 60 AA; 6933 MW; 0409C5311B453CF0 CRC64;
Query Match Score 22; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KRFK 4
Db 35 KRFK 36
RESULT 2
ID RS14_CLOPE STANDARD; PRT; 61 AA.
AC Q8XHC6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE 30S ribosomal protein S14.
GN RPSN OR CPB2392.
OC Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kihara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- FUNCTION: Known to be required for the assembly of 30S particles
CC and may also be responsible for determining the conformation of
CC the 16S rRNA at the A site (by similarity).
CC -1- SIMILARITY: Belongs to the S14P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AP003194; BAB82098.1; -
CC DR InterPro; IPR001209; Ribosomal_S14.
CC DR Pfam; PF00253; Ribosomal_S14; 1.
CC DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
CC KM Ribosomal protein; Complete proteome.
CC SQ SEQUENCE 61 AA; 7280 MW; C3910B3227493E75 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 35 KKF 38

RESULT 3
APM2 HUMAN STANDARD; PRT; 76 AA.
ID APM2_HUMAN
AC Q15847;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adipose most abundant gene transcript 2.
GN APM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=96224171; PubMed=8619847;
RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA Matsubara K.;
RT "cDNA cloning and expression of a novel adipose specific collagen-like
RT factor, apm1 (Adipose Most abundant Gene transcript 1).";
RL Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mallya S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Skalska U., Smalins D.E.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THE ADIPOSE TISSUE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; DA5370; BAA0826.1; -
CC DR EMBL; BC004471; AAH04471.1; -
CC SQ SEQUENCE 76 AA; 7855 MW; C95B41F13C4BA7B1 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 70 KKF 73

RESULT 4
R31B PHOLL STANDARD; PRT; 84 AA.
ID R31B_PHOLL
AC Q7NA98;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 50S ribosomal protein l31 type B.
GN RPEM2 OR RPEM OR PU00036.
OS Photobacterium luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photobacterus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Bournaud-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Médigue C., Lanois A., Powell K., Siguer P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photobacterus
RT luminescens";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -1- SIMILARITY: Belongs to the L31P family of ribosomal proteins.
CC Subfamily B.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; BX571859; CAB12331.1; -

DR Photolisc; plu0036; -

DR HAMAP; MF_00502; -; 1.

DR PROSITE; PS01143; RIBOSOMAL_L31; FALSE_NEG.

KM RIBOSOMAL protein; Complete proteome.

SO SEQUENCE 84 AA; 9593 MW; 39681B9E527105B2 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 75 KKF 78

RESULT 5

R31B STAM STANDARD; PRT; 84 AA.

AC 0995D8;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 50S ribosomal protein L31 type B.

GN RME2 OR RPME OR SAV2120 OR SA1922 OR MW2044.

OS Staphylococcus aureus (strain N315) / ATCC 700699).

OS Staphylococcus aureus (strain N315), and

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158878, 158879, 196620;

LN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Mu50 / ATCC 700699, and N315;

RX MEDLINE=21311952; PubMed=11418146;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T., Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-U. Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J., Kanetsu M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

RA "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."

RT Lancet 357:1225-1240(2001).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=Mu2;

RX MEDLINE=22040717; PubMed=12044378;

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A., Yamamoto K., Hiramatsu K.;

RA "Genome and virulence determinants of high virulence community-acquired MRSA."

RT Lancet 359:1819-1827(2002).

RL -1- SIMILARITY: Belongs to the L31P family of ribosomal proteins. Subfamily B.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AP00364; BAB58282.1; -

DR EMBL; AP003136; BAB43206.1; -

DR EMBL; AP004629; BAB95909.1; -

DR PIR; E90005; E90005.

DR HAMAP; MF_00502; -; 1.

DR InterPro; IPR002150; Ribosomal_L31.

DR Pfam; PF01197; Ribosomal_L31; 1.

DR PRINTS; PRO1249; RIBOSOMALL31.

DR TIGRFAMs; TIGR00105; L31; 1.

DR PROSITE; PS01143; RIBOSOMAL_L31; 1.

KM RIBOSOMAL protein; Complete proteome.

SO SEQUENCE 84 AA; 9723 MW; 28C86AE4FDEE01 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 76 KKF 79

RESULT 6

RL27_CAMO STANDARD; PRT; 84 AA.

AC 09FJ31;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 50S ribosomal protein L27.

GN RPWA OR CU0095.

OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;

OC Campylobacteraceae; Campylobacter.

OX NCBI_TaxID=197;

LN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 11168;

RX MEDLINE=20150912; PubMed=10688204;

RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C., Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S., Jorgensen K., Kariyasev A.V., Moule S., Pallen M.J., Penn C.W., Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M., Whitehead S., Barrall B.G.;

RA "The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences."

RT Nature 403:665-668(2000).

RL -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AL139074; CAB72579.1; -

DR PIR; G81425; G81425.

DR HAMAP; MF_00539; -; 1.

DR InterPro; IPR001684; Ribosomal_L27.

DR Pfam; PF01016; Ribosomal_L27; 1.

DR PRINTS; PRO0063; RIBOSOMAL_L27; 1.

DR PRODOM; PD003114; Ribosomal_L27; 1.

DR TIGRFAMs; TIGR00062; L27; 1.

DR PROSITE; PS00831; RIBOSOMAL_L27; 1.

KM RIBOSOMAL protein; Complete proteome.

SO SEQUENCE 84 AA; 9285 MW; EEC60824EB1B0670 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 24 KKF 27

```

RESULT 7
R13B_STAMP STANDARD; PRT; 85 AA.
ID R13B_STAMP
AC O8CRM8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L31 type B.
GN RMEW2 OR SE1718.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
ON NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Mao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
CC -i SIMILARITY: Belongs to the L31p family of ribosomal proteins.
CC Subfamily B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB016749; AA005317.1; -.
CC HAMAP; MF_00502; -; 1.
CC InterPro; IPR002150; Ribosomal_L31.
CC Pfam; PF01197; Ribosomal_L31; 1.
CC PRINTS; PR01249; RIBOSOMAL_L31.
CC TRGFAMS; TIGR00105; L31; 1.
CC DR TIGRPFAMS; PS01143; RIBOSOMAL_L31; 1.
CC DR PROSITE; PS01143; RIBOSOMAL_L31; 1.
CC KW Ribosomal protein; Complete proteome.
CC SQ SEQUENCE 85 AA; 9779 MW; 65B8D6687DCA5AAD CRC64;

Query Match 100.0%; Score 22; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFQ 4
DB 76 KKFQ 79

RESULT 8
R127_HELPJ STANDARD; PRT; 88 AA.
ID R127_HELPJ
AC Q9ZMD8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPYMA OR JHP0282.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
ON NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deLonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uita-Nickelsen M., Mills D.W., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;

```

```

RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -i SIMILARITY: Belongs to the L27p family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001465; AAD05863.1; -.
CC PIR; E71951; E71951.
CC HAMAP; MF_00539; -; 1.
CC InterPro; IPR001684; Ribosomal_L27.
CC Pfam; PF01016; Ribosomal_L27; 1.
CC PRINTS; PR00063; RIBOSOMAL_L27.
CC PRODOM; PD003114; Ribosomal_L27; 1.
CC TRGFAMS; TIGR00062; L27; 1.
CC DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
CC KW Ribosomal protein; Complete proteome.
CC SQ SEQUENCE 88 AA; 9676 MW; 29B0AE6341D73E60 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFQ 4
DB 24 KKFQ 27

RESULT 9
R127_HELPY STANDARD; PRT; 88 AA.
ID R127_HELPY
AC P56050;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPYMA OR HP0297.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
ON NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kierlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
CC -i SIMILARITY: Belongs to the L27p family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: A6000548; AAD07366.1; -
DR PIR: A64557; A64557.
DR TIGR: H0297; -
DR HAMAP: MF_00539; -; 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMALL27.
DR PRODOM: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
DR Riboosomal protein; Complete proteome.
KW SEQUENCE 88 AA; 9778 MW; 29B0AE61A7951F0 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KXFG 4
Db 24 KXFG 27

RESULT 10
RL27_SYN7 STANDARD; PRT; 88 AA.
AC 0923H6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPL27.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98397326; PubMed=9727980.
RA Ishiura M., Kutsuna S., Aoki S., Iwasaki H., Andersson C.R.,
RA Tanabe A., Golden S.S., Johnson C.H., Kondo T.;
RT "Expression of a gene cluster kaiABC as a circadian feedback process
in cyanobacteria";
RL Science 281:1519-1523(1998).

CC -1- SIMILARITY: Belongs to the L27 family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: A6010691; BA37099.1; -
DR PIR: T4265; T4265.
DR HAMAP: MF_00539; -; 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMALL27.
DR PRODOM: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
DR Riboosomal protein.
KW SEQUENCE 88 AA; 9226 MW; 0A448F019FFC42F0 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KXFG 4
Db 24 KXFG 27

RESULT 11
RL27_AGR5 STANDARD; PRT; 89 AA.
AC Q8UBR6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPL27 OR ATU2784 OR AGR_C_5052.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monke D.E., Kiteajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58";
RL Science 294:2317-2323(2001).

CC -1- SIMILARITY: Belongs to the L27 family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: A6009226; AAL4365.1; -
DR PIR: A608192; AAK8499.1; -
DR PIR: AG2918; AG2918.
DR HAMAP: MF_00539; -; 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMALL27.
DR PRODOM: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
DR Riboosomal protein; Complete proteome.
KW SEQUENCE 89 AA; 9396 MW; 59D141BF2B47301A CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KXFG 4
Db 24 KXFG 27

```

RESULT 12
RL27 BRUB STANDARD; PRT; 89 AA.
AC Q8W58;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L27.
GN RPLM OR RL27.
OS Brucella abortus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21952364; PubMed=11955619;
RA Halling S.M.; Zuercher R.L.;
RT "Evidence for lateral transfer to Brucellae: characterization of a
RT locus with a Tn-like element (Tn2020).";
RL Blochim Biophys Acta 1574:109-116(2002).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF119331; AAL3286.1; -
DR HAMAP; MF_00539; -; 1
DR InterPro; IPR001684; Ribosomal_L27.
DR Pfam; PF01016; Ribosomal_L27; 1.
DR PRINTS; PR00063; RIBOSOMAL_L27.
DR PRODOM; PD003114; Ribosomal_L27; 1.
DR TIGRFAMs; TIGR00062; L27; 1.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein.
SQ SEQUENCE 89 AA; 9377 MW; 8667B661EC97BAD5 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFQ 4
DB 24 KKFQ 27

RESULT 13
RL27 BRUM STANDARD; PRT; 89 AA.
AC Q8YJ84;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPLM OR BME10202.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G.; Kapatral V.; Redkar R.J.; Patra G.; Muier C.; Los T.;
RA Ivanova N.; Andersson I.; Bhattacharya A.; Lykidis A.; Reznik G.;
RA Jablonaki L.; Larsen N.; D'Souza M.; Bernal A.; Mazur M.; Goldsman E.;
RA Selkov E.; Elzer P.H.; Hagius S.; O'Callaghan D.; Teleson J.-J.;
RA Haseelkorn R.; Kyriades N.; Overbeek R.;

```

```

RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE009463; AAL5184.1; ALT_FRAME.
DR HAMAP; MF_00539; -; 1
DR InterPro; IPR001684; Ribosomal_L27.
DR Pfam; PF01016; Ribosomal_L27; 1.
DR PRINTS; PR00063; RIBOSOMAL_L27.
DR PRODOM; PD003114; Ribosomal_L27; 1.
DR TIGRFAMs; TIGR00062; L27; 1.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 89 AA; 9407 MW; 3D22A8838C9CAD09 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFQ 4
DB 24 KKFQ 27

RESULT 14
RL27 BRUSU STANDARD; PRT; 89 AA.
AC Q8FYB8;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 50S ribosomal protein L27.
GN RPLM OR BR1849.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T.; Seshadri R.; Nelson K.E.; Eisen J.A.; Heidelberg J.F.;
RA Read T.D.; Dodson R.J.; Umayam L.; Brinkac L.M.; Beanan M.J.;
RA Daugherty S.C.; Debroy R.T.; Durkin A.S.; Kolonay J.F.; Madupu R.;
RA Nelson W.C.; Ayodeji B.; Krahl M.; Shetty J.; Malek J.; Van Aken S.E.;
RA Tettmiller S.; Tettelin H.; Gill S.R.; White O.; Salzberg S.L.;
RA Hoover D.L.; Lindler L.E.; Halling S.M.; Boyle S.M.; Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014475; AAN30744.1; -
DR TIGR; BR1849; -

```

DR HAMAP: MF_00539; -: 1.
 DR InterPro: IPR001684; Ribosomal_L27.
 DR Pfam: PF01016; Ribosomal_L27; 1.
 DR PRINTS: PR00063; RIBOSOMAL_L27.
 DR ProDom: PD003114; Ribosomal_L27; 1.
 DR TIGRFAMs: TIGR00062; L27; 1.
 DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 89 AA; 9377 MW; 8667B661EC87BAD5 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFg 4
 ||||
 Db 24 KKFg 27

RESULT 15

RL27_CAUCR STANDARD; PRT; 89 AA.
 ID RL27_CAUCR
 AC O9ABB3;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 50S ribosomal protein L27.
 GN RPLM OR CC0318.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Ueberback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RL "Complete genome sequence of Caulobacter crescentus.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AE005705; AAK2305.1; -
 DR PIR: E87288; E87288.
 DR TIGR: CC0318; -
 DR HAMAP: MF_00539; -: 1.
 DR InterPro: IPR001684; Ribosomal_L27.
 DR Pfam: PF01016; Ribosomal_L27; 1.
 DR PRINTS: PR00063; RIBOSOMAL_L27.
 DR ProDom: PD003114; Ribosomal_L27; 1.
 DR TIGRFAMs: TIGR00062; L27; 1.
 DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 89 AA; 9371 MW; 22E85AECF5A28C91 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFg 4

Db 24 KKFg 27

Search completed: March 18, 2004, 11:17:00
 Job time: 1.24508 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 1.09409 Seconds
(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-3
Perfect score: 22
Sequence: 1 KKG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	25	10	094199 atropa bell
2	22	100.0	30	16	098MC2 rhizobium I
3	22	100.0	35	15	097875 human immun
4	22	100.0	50	16	097732 streptococ
5	22	100.0	50	16	08DRH5 streptococ
6	22	100.0	58	16	09KPA9 streptococ
7	22	100.0	59	5	096380 aedes albop
8	22	100.0	59	5	086PR6 culic pipie
9	22	100.0	59	16	0831G6 enterococcu
10	22	100.0	60	11	035588 mesocricetu
11	22	100.0	60	17	082Z07 pyrobaculum
12	22	100.0	61	16	08U4W9 agrobacteri
13	22	100.0	61	16	07VCU4 prochloroco
14	22	100.0	62	10	09L0G1 arabidopsis
15	22	100.0	62	16	P7618 escherichia
16	22	100.0	63	16	08DL07 synechococc

17	22	100.0	64	10	09FRT4	09fnc4 brassica ol
18	22	100.0	65	16	08DB17	08db17 vibrio vuln
19	22	100.0	65	16	081TK0	081tk0 bacillus an
20	22	100.0	66	9	085E63	085e63 mycobacteri
21	22	100.0	66	10	024089	024089 medicago tr
22	22	100.0	66	16	081LX1	081lx1 bacillus an
23	22	100.0	66	16	0812T5	0812t5 bacillus ce
24	22	100.0	67	12	091BL0	091bl0 spodoptera
25	22	100.0	67	13	093340	093340 struthio ca
26	22	100.0	68	16	087GC6	087gc6 vibrio para
27	22	100.0	69	16	08B6P6	08b6p6 streptococ
28	22	100.0	69	16	08B191	08b191 streptococ
29	22	100.0	69	16	087HL5	087hl5 vibrio para
30	22	100.0	69	16	081GFI	081gfi bacillus ce
31	22	100.0	70	10	09SXM3	09sxm3 lithospermu
32	22	100.0	70	10	09XHT1	09xht1 brassica ol
33	22	100.0	72	17	0971I4	0971i4 sulfolobus
34	22	100.0	73	16	088WQ1	088wq1 lactobacilli
35	22	100.0	76	5	08T6A2	08t6a2 aplysia cal
36	22	100.0	76	15	089369	089369 human immun
37	22	100.0	76	16	08RHG0	08rhg0 fusobacteri
38	22	100.0	77	2	08B142	08b142 vibrio para
39	22	100.0	77	10	084JQ2	084jq2 arabidopsis
40	22	100.0	78	5	086W41	086w41 phylliopterh
41	22	100.0	78	11	091Y28	091y28 mus musculu
42	22	100.0	81	15	068996	068996 human endog
43	22	100.0	81	16	07UPR5	07upr5 rhodospirell
44	22	100.0	82	16	08G4U1	08g4u1 bifidobacte
45	22	100.0	83	11	062306	062306 mus musculu

ALIGNMENTS

RESULT 1

ID 094199 PRELIMINARY; PRT; 25 AA.

AC 094199;
DT 01-DEC-2001 (T-REMBLrel. 19, Created)
DT 01-DEC-2001 (T-REMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (T-REMBLrel. 24, Last annotation update)
DE 40S ribosomal protein (Fragment).
OS Atropa belladonna (Belladonna) (Deadly nightshade).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Atropa.
OX NCBI_TaxID=31113;
RN [1]
RP SEQUENCE FROM N.A.
RA Nouar E., Baucher M., Jaziri M.;
RT "Differential gene expression in Atropa belladonna leafy gall.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA El Hassan N.;
RL Thesis (2001), Department of Plant Biotechnology,
RL Universite Libre De Bruxelles, Bruxelles, Belgium.
DR EMBL: AJ291783; CAC44248.1; --
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0005840; C:ribosome; IEA.
DR GO: GO:0003735; P:structural constituent of ribosome; IEA.
DR GO: GO:006412; P:protein biosynthesis; IEA.
DR InterPro: IPR000754; Ribosomal_S9.
DR Pfam: PF00380; Ribosomal_S9; 1.
DR PfDom: PD001627; Ribosomal_S9; 1.
KW Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 25 AA; 2986 MW; BFF06F7182304F69 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 9 KKF 12

RESULT 2

O98MC2 PRELIMINARY; PRT; 30 AA.
 AC O98MC2;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein ms10637.
 GN MS10637.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneke T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP002995; BAB48191.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 30 AA; 3302 MW; 9DDIDA28C11C657 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 30;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 2 KKF 5

RESULT 3

ID O97875 PRELIMINARY; PRT; 35 AA.
 AC O97875;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Gp160 protein (Fragment).
 GN ENV OR ENV, V3.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VI526;
 RX MEDLINE=96426454; PubMed=8828748;
 RA Delaporte E., Janssens W., Peeters M., Buve A., Dhangra G.,
 RA Perret J.L., Ditsambou V., Georges Courtbot M.C., Georges A.,
 RA Bougeois A., Samb B., Henzel D., Heyndrickx L., Franssen K.,
 RA Van der Groen G., Larouz B., Mbe J.R.;
 RT "Epidemiological and Molecular characteristics of HIV infection in
 Gabon (1986 - 1994)."
 RL AIDS 10:903-910(1996).
 DR EMBL; X90922; CAA62429.1;
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 KW AIDS; Coat protein; Glycoprotein.

FT NON TER 1 1
 PT NON TER 35 35
 SQ SEQUENCE 35 AA; 3847 MW; 81AF1F7D04503244 CRC64;

Query Match

Best Local Similarity 100.0%; Score 22; DB 15; Length 35;
 Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 12 KKF 15

RESULT 4

O97732 PRELIMINARY; PRT; 50 AA.
 ID O97732;
 AC O97732;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein SP0124.
 GN SP0124.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC BAA-334 / TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzaple E., Khouri H., Wolf A.M., Ueberback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 pneumoniae."
 RL Science 293:498-506(2001).
 DR EMBL; AE007329; AAK74310.1; -.
 DR PIR; E95014; E95014.
 DR TIGR; SP0124; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 50 AA; 5558 MW; 9393EB8E261530BF CRC64;

Query Match 100.0%; Score 22; DB 16; Length 50;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 44 KKF 47

RESULT 5

ID O98RH5 PRELIMINARY; PRT; 50 AA.
 AC O98RH5;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Orf51.
 GN ORF51 OR SP0127.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,

RA Dehoff B.S., Beltram S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.B.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., Mcleaster K., Mundy C.W., Nicot T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun F.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaekunas S.R., Rostock P.R., Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL: AEO08396; AAK98931.1; --
 DR PIR: G97887; G97887.
 KW Complete proteome.
 SQ SEQUENCE 50 AA; 5557 MW; 7E2888EE26153897 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 50;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
 Db 44 KKFg 47

RESULT 6

Q9KFA9 PRELIMINARY; PRT; 58 AA.
 AC Q9KFA9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein BH0576.
 GN BH0576.
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=8665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001509; BAB04295.1; --
 DR PIR: H83721; H83721.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 58 AA; 6353 MW; 40D50F0E76BD5E83 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 58;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
 Db 30 KKFg 33

RESULT 7

Q963B0 PRELIMINARY; PRT; 59 AA.
 AC Q963B0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Antiheliotic peptide cecropin A2.
 GN CECRA2.
 OS *Aedes albopictus* (Forest day mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 OX NCBI_TaxID=7160;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Sun D., Fallon A.;
 RT "Characterization of genomic DNA encoding mosquito cecropins.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF394745; AAK81850.1; --
 SQ SEQUENCE 59 AA; 6183 MW; C5C1DFE3D8A8673 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
 Db 27 KKFg 30

RESULT 8

Q86PR6 PRELIMINARY; PRT; 59 AA.
 AC Q86PR6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cecropin A.
 OS *Culex pipiens pipiens* (Northern house mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culex.
 OX NCBI_TaxID=38569;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Iowa State;
 RA Bartholomay L.C., Farid H.A., Ramzy R.M., Christensen B.M.;
 RT "Immune immunity in the *Culex pipiens*-*Wuchereria bancrofti* host-
 RT parasite relationship.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY189808; AAO38516.1; --
 SQ SEQUENCE 59 AA; 6285 MW; 26CE3D7E823D9296 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
 Db 27 KKFg 30

RESULT 9

Q831G6 PRELIMINARY; PRT; 59 AA.
 AC Q831G6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Transcriptional regulator, putative.
 GN EF2543.
 OS *Enterococcus faecalis* (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
 RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
 RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Kouri H.,
 RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 RL *Enterococcus faecalis*.";
 Science 299:2071-2074(2003).

DR EMBL; AE016955; AAC82256.1; --
 DR TIGR; EF2543; --
 KM Complete proteome.
 SQ SEQUENCE 59 AA; 6731 MW; FCC4A51EF3BCAD24 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 DB 52 KKF 55

RESULT 10
 035588 PRELIMINARY; PRT; 60 AA.
 AC 035588;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE C11 protein (Fragment).
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 NC NCBI_TaxID=10036;
 RX MEDLINE=95082951; PubMed=7990965;
 RA Kress M., Arman I., Haenni A.L., Celis J.E., Philippe M., Justesen K.,
 RA Kisseliev L.,
 RA "A highly conserved eukaryotic protein family possessing properties of
 RT polypeptide chain release factor."
 RL Nature 372:701-703(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Baby Kidney;
 RX MEDLINE=97338483; PubMed=9195043;
 RA Uebera B., Ewilaichitr L., Stanfield I., Taasan J.P., Le Goff X.,
 RA Kress M., Tuite M.F.,
 RA "Expression of the release factor eRF1 (Sup45p) gene of higher
 RT eukaryotes in yeast and mammalian tissues."
 RL Biochimie 79:27-36(1997).
 DR EMBL; X81624; CAAS7280.1; --
 DR HSSP; P46055; 1DT9.
 DR InterPro; IPR005142; eRF1_3.
 DR Pfam; PF03465; eRF1_3; 1.
 FT NON TER 1
 SQ SEQUENCE 60 AA; 6785 MW; 57F075EA624FEF6 CRC64;

Query Match 100.0%; Score 22; DB 11; Length 60;
 Best Local Similarity 100.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 DB 6 KKF 9

RESULT 11
 082207 PRELIMINARY; PRT; 60 AA.
 AC 082207;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein PA0507.
 GN PA0507.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;

OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX MEDLINE=21664397; PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.,
 RA "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL; AE009771; AAL62834.1; --
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 60 AA; 7290 MW; 89171FD9473B8AD2 CRC64;

Query Match 100.0%; Score 22; DB 17; Length 60;
 Best Local Similarity 100.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 DB 20 KKF 23

RESULT 12
 08U4W9 PRELIMINARY; PRT; 61 AA.
 AC 08U4W9;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE AGR_L_2351P.
 GN AGR_L_2351.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NC NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Galtung S., Miller N., Blanchard M.,
 RA Quirillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmel K., Gordon M., Vaudin M., Tarchouk O., Epp A., Liu F.,
 RA Wolam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Guxson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RA "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL; AE008318; AAK89742.1; --
 DR PIR; D98277; D98277.
 SQ SEQUENCE 61 AA; 6772 MW; 61SDA6FA3A0A56D2 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 61;
 Best Local Similarity 100.0%; Pred. No. 6.5e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 DB 32 KKF 35

RESULT 13
 07VCU4 PRELIMINARY; PRT; 61 AA.
 AC 07VCU4;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Predicted protein family PM-6.
 GN PRO0646.
 OS Prochlorococcus marinus.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcales;

OC Prochlorococcus.
 OX NCBI_TaxID=1219;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SARG / CCM 1375 / SS120;
 RX MEDLINE=22810154; PubMed=12917466;
 RA Dufrenoy A., Salanoubat M., Fartensky F., Artiguenave F., Axmann I.M.,
 Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
 RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
 RA Scanlan D.J., Tandeau de Marsac N., Welzenbach J., Wincker P.,
 RA Wolf Y.I., Hess W.R.;
 RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
 RT a nearly minimal oxygenotrophic genome.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
 DR EMBL; AB017162; AAP9690.1; -;
 KM Complete proteome.
 SQ SEQUENCE 61 AA; 6747 MW; AC032EF5BF7997B CRC64;

Query Match 100.0%; Score 22; DB 16; Length 61;
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKEG 4
 DB 12 KKEG 15.

RESULT 14

O9LOG1 PRELIMINARY; PRT; 62 AA.

AC O9LOG1;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE P1504.24.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Becker J.R.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Becker J.R.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooke S., Buehler E., Chao O., Johnson-Hopson C.,
 RA Khan S., Kim C., Altati H., Bel B., Chin C., Chieu J., Choi E., Lam B.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharbek N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Becker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007887; AAF79552.1; -;
 SQ SEQUENCE 62 AA; 7607 MW; DD8DF41F1624FEC4 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 62;
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKEG 4
 DB 24 KKEG 27

RESULT 15
 P76618 PRELIMINARY; PRT; 62 AA.

AC P76618;
 DT 01-FEB-1997 (TEMBLrel. 02, Created)
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE From BARS 275730 to 2786129 (Section 240 of 400) of the complete
 DE genome (Section 240 of 400).
 GN B2656.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.O.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 DR EMBL; AE000350; AAC75703.1; -;
 DR PIR; A65045; A65045.
 KM Complete proteome.
 SQ SEQUENCE 62 AA; 6714 MW; ED2C3CC9B87C20A2 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 62;
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKEG 4
 DB 15 KKEG 18

Search completed: March 18, 2004, 11:19:19
 Job time : 3.09409 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 4.9628 Seconds
(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-4

Perfect score: 67

Sequence: 1 ASTDTWGRPLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	12	5	AAG79462 sc-uPA 10
2	67	100.0	86	7	ABR42617 Human abr
3	67	100.0	86	7	ABR42602 Human abr
4	67	100.0	86	7	ABR42599 Human abr
5	67	100.0	86	7	ABR42601 Human abr
6	67	100.0	87	7	ABR42604 Human abr
7	67	100.0	87	7	ABR42605 Human abr
8	67	100.0	88	5	AAE16542 Human uro
9	67	100.0	89	2	AAW22742 Urokinase
10	67	100.0	96	5	AAE16550 Human uPA
11	67	100.0	100	1	AAE80996 Engineere
12	67	100.0	135	5	AAE16545 Human uro
13	67	100.0	143	5	AAE16549 Human uPA
14	67	100.0	201	2	AAW22746 Metastasi
15	67	100.0	208	2	AAW22747 Metastasi
16	67	100.0	322	7	ABR42615 Fusion pr
17	67	100.0	322	7	ABR42616 Fusion pr
18	67	100.0	337	4	ABG75492 Human col
19	67	100.0	337	5	ABP41795 Human ova
20	67	100.0	365	2	AAE68854 Delta 1-4
21	67	100.0	378	2	AAW13635 Human pro
22	67	100.0	386	2	AAE66266 Bifunctio
23	67	100.0	389	2	AAW13636 Human pro
24	67	100.0	390	2	AAE66245 Bifunctio
25	67	100.0	390	2	AAE66247 Bifunctio

26	67	100.0	392	2	AAE66260 Bifunctio
27	67	100.0	392	2	AAE66264 Bifunctio
28	67	100.0	392	2	AAE66255 Bifunctio
29	67	100.0	392	2	AAE66259 Bifunctio
30	67	100.0	392	2	AAE66261 Bifunctio
31	67	100.0	392	2	AAE66258 Bifunctio
32	67	100.0	392	2	AAE66263 Bifunctio
33	67	100.0	392	2	AAE66254 Bifunctio
34	67	100.0	392	2	AAE66256 Bifunctio
35	67	100.0	392	2	AAE66257 Bifunctio
36	67	100.0	393	2	AAE66251 Bifunctio
37	67	100.0	393	2	AAE66253 Bifunctio
38	67	100.0	393	2	AAE66249 Bifunctio
39	67	100.0	393	2	AAE66252 Bifunctio
40	67	100.0	393	2	AAE66244 Bifunctio
41	67	100.0	393	2	AAE66250 Bifunctio
42	67	100.0	393	2	AAE66257 Bifunctio
43	67	100.0	393	2	AAE66256 Bifunctio
44	67	100.0	393	2	AAE66255 Bifunctio
45	67	100.0	395	2	AAE66247 Bifunctio

ALIGNMENTS

RESULT 1
AAG79462 standard; peptide; 12 AA.
ID AAG79462:
AC AAG79462:
AC AAG79462:
DT 15-NOV-2002 (first entry)
XX
XX
DE sc-uPA long A chain peptide fragment 2, amino acids 82-93.
XX
XX
KW Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
KW high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KW long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KM low molecular weight urokinase-type plasminogen activator; LMW-uPA; CD87;
KM binding domain.
XX
XX Homo sapiens.
OS
XX
XX
PN EP1232755-A2.
XX
XX
PD 21-AUG-2002..
XX
XX
PF 15-FEB-2002; 2002EP-00003555.
XX
XX
PR 20-FEB-2001; 2001JP-00042655.
PR 19-JUN-2001; 2001JP-00184284.
XX
XX (JGRP-) JCR PHARM CO LTD.
XX
XX Wada M, Wada N;
PI
XX
XX WPI; 2002-610512/66.
DR
XX
XX
PT Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g. high
PT molecular weight urokinase-type plasminogen activator, amino-terminal
PT fragment or an anti-CD87 antibody.
XX
XX
PS Example; Page 23; 38pp; English.
XX
XX The sequences given in AAG79461-63 represent peptide fragments derived
XX from the long A chain of single chain prepro-urokinase (sc-uPA). Pro-
XX urokinase (amino acids 21-431) with a cleavage between amino acids 178
XX and 179 gives high molecular weight urokinase-type plasminogen activator
XX (HMW-uPA). HMW-uPA is a protein consisting of two peptide chains linked
XX by a disulphide bond. The chains, long A and B, are formed by enzymatic
XX cleavage between amino acids 178 and 179 of pro-urokinase. HMW-uPA
XX includes an EGF-like domain, a kringle domain and a urokinase receptor
XX (CD87) binding domain. HMW-uPA is then cleaved between amino acids 155

[illegible]

FT	Domain	54..60	/note="kringle domain"
XX			
XX	WO2003042354-A2.		
XX			
PD	22-MAY-2003.		
XX			
XX	04-SEP-2002; 2002WO-US027885.		
PF			
XX	04-SEP-2001; 2001US-0316300P.		
PR			
XX			
XX	(AVET) AVENTIS PHARM INC.		
PA			
XX			
PI	Nesbit M, Fong TC, Brockstedt D;		
XX			
XX	WPI; 2003-449566/42.		
DR			
XX			
XX	New abrogen polypeptide, useful for treating an angiogenesis related		
PT	diseases e.g. tumor metastasis.		
XX			
PS	Claim 37; Page 26; 95pp; English.		
XX			
XX	The present sequence is the protein sequence of a secreted form of a		
CC	novel human abrogen, designated hATP-kringle, comprising the human		
CC	urokinase plasminogen activator kringle domain. The sequence includes an		
CC	N-terminal alanine residue that results from cleavage of an interleukin-2		
CC	signal peptide. Abrogens such as hATP-kringle are potent inhibitors of		
CC	endothelial proliferation and angiogenesis. Abrogen polypeptides are		
CC	capable of inhibiting or reducing cell proliferation induced by both		
CC	basic fibroblast growth factor (bFGF) and vascular endothelial growth		
CC	factor in a specific endothelial cell proliferation assay; angiotatin		
CC	only inhibits bFGF induced proliferation in this assay. Vectors that		
CC	expressed abrogen polypeptides in vivo were shown to reduce tumour		
CC	metastasis in 2 lung cancer models. The invention provides abrogen		
CC	polypeptides and polynucleotides, and methods of using these to treat an		
CC	angiogenesis-related disease or disorder, e.g. tumour metastasis		
CC	(claimed)		
XX			
XX	Sequence 87 AA;		
XX			
XX	Query Match	100.0%; Score 67; DB 7; Length 87;	
XX	Best Local Similarity	100.0%; Pred. No. 0.00074;	
XX	Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ASTDTMGRPCLP 12		
DB	16 ASTDTMGRPCLP 27		
XX			
XX	RESULT 8		
XX	AAE16542		
ID	AAE16542 standard; protein; 88 AA.		
XX			
XX	AAE16542;		
XX			
DT	09-APR-2002 (first entry)		
XX			
DE	Human urokinase-type plasminogen activator (uPA) kringle.		
XX			
XX	Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;		
XX	stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;		
KW	microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;		
KW	tumour cell metastasis; glioma; diabetic retinopathy; wound healing;		
KW	clotting disorder; uterine contraction disorder; respiratory disease;		
KW	male impotence; adult respiratory distress syndrome.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200197752-A2.		
XX			
PD	27-DEC-2001.		
XX			
XX	13-JUN-2001; 2001WO-US018976.		

XX 20-JUN-2000; 2000US-0212874P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX Cines DB, Higazi AA;
 XX WPI; 2002-122240/16.
 DR N-PSDB; AAD27075.
 XX
 PT Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 XX
 PS Claim 1; Fig 1A; 117pp; English.
 XX
 CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle
 CC
 XX Sequence 88 AA;

Query Match 100.0%; Score 67; DB 5; Length 88;
 Best Local Similarity 100.0%; Pred. No. 0.00075; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 0;

QY 1 ASTDTMGRPCLP 12
 |||||
 DB 15 ASTDTMGRPCLP 26

RESULT 9
 AAW22742
 ID AAW22742 standard; protein; 89 AA.
 XX AAW22742;

AC 12-MAR-1998 (first entry)
 XX
 DT Urokinase residues 43-131.
 XX

KW Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
 KM G-domain; urokinase; prevention; leukaemia; lymphoma.
 XX

OS Homo sapiens.
 XX
 PN MO9725422-A1.
 XX

PD 17-JUL-1997.
 XX

PF 06-JAN-1997; 97MO-JP000008.
 XX

PR 08-JAN-1996; 96JP-00001059.
 XX

PA (NISP) NISSIN FOOD PROD CO LTD.
 XX

PI Kobayashi H, Terao T, Sugino D, Okushima M;
 XX

DR WPI; 1997-372862/34.
 XX

PT Chimeric protease which inhibits development of metastases in cancer -
 PT contains urinary trypsin inhibitor carboxy-terminal domain linked to

PT urokinase G-domain.
 XX
 PS Claim 3; Page 72; 97pp; Japanese.
 XX

CC A novel chimeric protein contains the carboxy-terminal domain of human
 CC urinary trypsin inhibitor (HI-8), which inhibits cancer cell metastasis,
 CC linked to a peptide containing the G-domain of urokinase (AAW22742),
 CC which specifically binds the excess urokinase receptor expressed in
 CC cancer cells. The chimeric protein has the amino-terminal AAW22734, the
 CC carboxy-terminal AAW22735 and a linking sequence selected from AAW22736-
 CC 39 or partial sequences derived from these, specifically AAW38130-63. The
 CC chimeric protein may also have additional amino-terminal sequences
 CC selected from AAW22740 or 9 partial sequences derived from this, and/or
 CC additional carboxy-terminal sequences selected from AAW22743 or 10
 CC partial sequences derived from this. The chimeric protein can be used to
 CC prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
 CC stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
 CC prostate, and in leukaemia or lymphoma
 CC
 XX Sequence 89 AA;

Query Match 100.0%; Score 67; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 0.00076; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 0;

QY 1 ASTDTMGRPCLP 12
 |||||
 DB 20 ASTDTMGRPCLP 31

RESULT 10
 AAE16550
 ID AAE16550 standard; protein; 96 AA.
 XX AAE16550;

AC 09-APR-2002 (first entry)
 XX

DE Human uPA kringle and connecting peptide.
 XX

KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; male impotence.
 XX

OS Homo sapiens.
 XX
 PN MO200197752-A2.
 XX

PD 27-DEC-2001.
 XX

PF 13-JUN-2001; 2001WO-US018976.
 XX

PR 20-JUN-2000; 2000US-0212874P.
 XX

PA (UYPE-) UNIV PENNSYLVANIA.
 XX

PI Cines DB, Higazi AA;
 XX

DR WPI; 2002-122240/16.
 XX

PR N-PSDB; AAD27083.
 XX

PT Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 XX

PS Claim 25; Fig 11; 117pp; English.
 XX

CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,

endothelial cell or tissue. The composition is used for treating stroke, hypertension, hyperextension, atherosclerosis, heart attack, microvascular occlusions, thrombotic microangiopathies, surgically induced thrombotic disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is human urokinase-type plasminogen activator (uPA) kringle and connecting peptide

CC Sequence 96 AA;

Query Match 100.0%; Score 67; DB 5; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.00082;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
DB 15 ASTDTMGRPCLP 26

RESULT 11

AA080996 ID A080996 standard; protein; 100 AA.

AC A080996;

AC 25-MAR-2003 (revised)

DT 08-OCT-1990 (first entry)

DE Engineered Human prourokinase mutant.

XX human pro-urokinase mutant; fibrinolysis; epidermal growth factor domain;
KW ss.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 70 /label= mutation
FT /note= "Wild-type= Cys; Mutant= Ser."

XX EP253241-A.

XX 20-JAN-1988.

XX 03-JUN-1987; 87EP-00109628.

XX 03-JUN-1986; 86JP-00156936.

XX 18-FEB-1987; 87JP-00036495.

XX (GREC) GREEN CROSS CORP.

XX Kasei S, Hiramatsu R, Uno S, Nagai M, Arimura H;

XX WPI; 1988-015623/03.

XX N-PSDB; AAN81299.

XX New human pro-urokinase mutants with fibrinolytic activity - have longer
XX half-lives in blood and are obtained by recombinant DNA procedures.

XX Disclosure; Page 7; 40p; English.

XX The DNA encoding this peptide contains two mutations introduced by site -
XX directed mutagenesis to create two unique restriction sites; the first is
XX recognised by SacI and the second by NdeI. The mutant amino acid denoted
XX in the features is encoded by the first 3 bases of the NdeI site. The
XX invention covers three deletion mutants derived from pro-kinase. All
XX three begin from amino acid 21 (Ser) and have either 1) Asn(30) to
XX Glu(63) and Thr(69) to Ser(70); or 2) Asn(30) to Lys(66) and Ser(70); or

CC 3) Asn(30) to Tyr(71) deleted. The deleted region contains the Epidermal
CC Growth Factor domain of the enzyme. (Updated on 25-MAR-2003 to correct PR
CC field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 100 AA;

Query Match 100.0%; Score 67; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
DB 82 ASTDTMGRPCLP 93

RESULT 12

AA016545 ID A016545 standard; protein; 135 AA.

AC A016545;

AC 09-APR-2002 (first entry)

DE Human urokinase-type plasminogen activator amino terminal fragment (ATF).

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX stroke; hypertension; atherosclerosis; heart attack; thrombotic disorder;
XX microvascular occlusion; angiotensin disorder; pulmonary fibrosis; asthma;
XX tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
XX clotting disorder; uterine contraction disorder; respiratory disease;
XX adult respiratory distress syndrome; amino terminal fragment; ATF;
XX male impotence.

XX Homo sapiens.

XX W020019752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US018976.

XX 20-JUN-2000; 2000US-0212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

XX N-PSDB; AAD27078.

XX Composition for modulating muscle cell and tissue contractility for
XX treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
XX comprising domains from urokinase-type plasminogen activator.

XX Claim 11; Fig 1D; 117pp; English.

XX The invention relates to a composition comprising one or more domains of
XX urokinase-type plasminogen activator (uPA). The composition is used to
XX modulate the contractility and angiogenic activity of a mammalian muscle,
XX endothelial cell or tissue. The composition is used for treating stroke,
XX hypertension, hyperextension, atherosclerosis, heart attack, microvascular
XX occlusions, thrombotic microangiopathies, surgically induced thrombotic
XX disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell
XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
XX diabetic retinopathy, wound healing, clotting disorder, uterine
XX contraction disorder, male impotence, respiratory disease or condition
XX such as asthma, adult respiratory distress syndrome, primary pulmonary
XX hypertension, microvascular thrombotic occlusion, and a disorder
XX associated with chronic intrapulmonary fibrin formation. The present
XX sequence is human urokinase-type plasminogen activator (uPA) amino
XX terminal fragment (ATF)

XX Sequence 135 AA;

Query Match 100.0%; Score 67; DB 5; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
DB 62 ASTDTMGRPCLP 73

RESULT 13

ID AAE16549 standard; protein; 143 AA.

AAE16549;

09-APR-2002 (first entry)

Human uPA amino terminal fragment (ATF) and connecting peptide.

Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
clotting disorder; uterine contraction disorder; respiratory disease;
adult respiratory distress syndrome; amino terminal fragment; ATF;
male impotence.

Homo sapiens.

MO200197752-A2.

27-DEC-2001.

13-JUN-2001; 2001MO-US018976.

20-JUN-2000; 2000US-0212874P.

(UYPE-) UNIV PENNSYLVANIA.

Cines DB, Higazi AA;

WPI; 2002-122240/16.

N-PSDB; AAD27082.

Composition for modulating muscle cell and tissue contractility for
treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
comprising domains from urokinase-type plasminogen activator.

Claim 24; Fig 1H; 117pp; English.

The invention relates to a composition comprising one or more domains of
urokinase-type plasminogen activator (uPA). The composition is used to
modulate the contractility and angiogenic activity of a mammalian muscle,
endothelial cell or tissue. The composition is used for treating stroke,
hypertension, atherosclerosis, heart attack, microvascular
occlusions, thrombotic microangiopathies, surgically induced thrombotic
disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
diabetic retinopathy, wound healing, clotting disorder, uterine
contraction disorder, male impotence, respiratory disease or condition
such as asthma, adult respiratory distress syndrome, primary pulmonary
hypertension, microvascular thrombotic occlusion, and a disorder
associated with chronic interpulmonary fibrin formation. The present
sequence is human urokinase-type plasminogen activator (uPA) amino
terminal fragment (ATF) and connecting peptide

Sequence 143 AA;

Query Match 100.0%; Score 67; DB 5; Length 143;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
DB 62 ASTDTMGRPCLP 73

RESULT 14

ID AAW22746 standard; protein; 201 AA.

AAW22746;

12-MAR-1998 (first entry)

Metastasis inhibitor.

Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
G-domain; urokinase; prevention; leukaemia; lymphoma.

Homo sapiens.

MO9725422-A1.

17-JUL-1997.

06-JAN-1997; 97MO-JP000008.

08-JAN-1996; 96JP-00001059.

(NISP) NISSIN FOOD PROD CO LTD.

Kobayashi H, Terao T, Sugino D, Okushima M;

WPI; 1997-372862/34.

N-PSDB; AAT75154.

Chimeric protein which inhibits development of metastases in cancer -
contains urinary trypsin inhibitor carboxy-terminal domain linked to
urokinase G-domain.

Claim 18; Page 57-58; 97pp; Japanese.

The present sequence is a novel chimeric protein, which contains the
carboxy-terminal domain of human urinary trypsin inhibitor (HI-8), which
inhibits cancer cell metastasis, linked to a peptide containing the G-
domain of urokinase, which specifically binds the excess urokinase
receptor expressed in cancer cells. The chimeric protein can be used to
prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
prostate, and in leukaemia or lymphoma

Sequence 201 AA;

Query Match 100.0%; Score 67; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
DB 63 ASTDTMGRPCLP 74

RESULT 15

ID AAW22747 standard; protein; 208 AA.

AAW22747;

12-MAR-1998 (first entry)

Metastasis inhibitor.

Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
G-domain; urokinase; prevention; leukaemia; lymphoma.

XX Homo sapiens.
 OS
 XX
 PN WO9725422-A1.
 XX
 PD 17-JUL-1997.
 XX
 PF 06-JAN-1997; 97WO-JP000008.
 XX
 PR 08-JAN-1996; 96JP-00001059.
 XX
 PA (NISP) NISSIN FOOD PROD CO LTD.
 XX
 PI Kobayashi H, Terao T, Sugino D, Okushima M;
 DR WPI; 1997-372862/34.
 DR N-PSDB; AAT75155.
 XX
 PT Chimeric protein which inhibits development of metastases in cancer -
 PT contains urinary trypsin inhibitor carboxy-terminal domain linked to
 PT urokinase G-domain.
 XX
 PS Claim 19; Page 59-60; 97pp; Japanese.
 XX
 CC The present sequence is a novel chimeric protein, which contains the
 CC carboxy-terminal domain of human urinary trypsin inhibitor (HI-8) which
 CC inhibits cancer cell metastasis, linked to a peptide containing the G-
 CC domain of urokinase, which specifically binds the excess urokinase
 CC receptor expressed in cancer cells. The chimeric protein can be used to
 CC prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
 CC stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
 CC prostate, and in leukemia or lymphoma
 XX
 SQ Sequence 208 AA;

Query Match 100.0%; Score 67; DB 2; Length 208;
 Best Local Similarity 100.0%; Pred. No. 0.0018;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASTDTWGRPCLP 12
 |||||
 Db 63 ASTDTWGRPCLP 74

Search completed: March 18, 2004, 11:16:23
 Job time : 5.9628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:13:05 ; Search time 1.39168 Seconds
(without alignments)
445.152 Million cell updates/sec

Title: US-10-076-421-4

Perfect score: 67

Sequence: 1 ASTDTMGRPCLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB pep.*
5: /cgn2_6/ptodata/2/1aa/PTCUG COMB pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	89	4	US-09-101-272G-62
2	67	100.0	138	2	US-08-797-689-12
3	67	100.0	138	4	US-09-984-186-12
4	67	100.0	157	3	US-08-142-5908-25
5	67	100.0	194	4	US-09-101-272G-80
6	67	100.0	200	4	US-09-101-272G-73
7	67	100.0	201	4	US-09-101-272G-96
8	67	100.0	208	4	US-09-101-272G-98
9	67	100.0	365	1	US-08-093-741-83
10	67	100.0	365	1	US-08-720-012-83
11	67	100.0	393	2	US-08-560-098A-44
12	67	100.0	393	3	US-08-967-024C-24
13	67	100.0	393	3	US-08-967-024C-25
14	67	100.0	411	1	US-08-087-163-1
15	67	100.0	411	1	US-08-286-7488-18
16	67	100.0	411	1	US-08-153-799-18
17	67	100.0	411	2	US-08-560-098A-48
18	67	100.0	411	3	US-09-181-816-1
19	67	100.0	411	4	US-09-403-736-2
20	67	100.0	430	1	US-07-942-157A-3
21	67	100.0	430	6	5219569-2
22	67	100.0	431	4	US-09-101-272G-1
23	67	100.0	431	6	5188829-1
24	67	100.0	432	1	US-08-560-098A-47
25	40	59.7	155	1	US-08-051-142-2
26	38	56.7	151	4	US-09-489-039A-11146
27	38	56.7	591	2	US-08-468-249A-20

28	38	56.7	593	2	US-08-468-249A-21	Sequence 21, App1
29	37	55.2	288	4	US-09-252-991A-29045	Sequence 29045, A
30	36	53.7	188	4	US-09-134-000C-4574	Sequence 4574, Ap
31	36	53.7	160	4	US-09-800-729-191	Sequence 191, App
32	36	53.7	541	4	US-08-311-731A-171	Sequence 171, App
33	36	53.7	556	3	US-08-687-590-24	Sequence 24, App1
34	35.5	53.0	118	4	US-09-252-991A-32474	Sequence 32474, A
35	35	52.2	15	4	US-09-252-991A-27145	Sequence 27145, A
36	35	52.2	137	4	US-09-489-039A-8628	Sequence 8628, Ap
37	35	52.2	202	4	US-09-540-236-3054	Sequence 3054, Ap
38	35	52.2	399	4	US-09-252-991A-25295	Sequence 25295, A
39	35	52.2	433	3	US-08-691-563C-87	Sequence 87, App1
40	35	52.2	433	4	US-09-374-766-87	Sequence 87, App1
41	35	52.2	433	4	US-08-979-847B-81	Sequence 81, App1
42	35	52.2	525	4	US-09-252-991A-20491	Sequence 20491, A
43	35	52.2	787	4	US-09-721-183-2	Sequence 2, App1
44	35	52.2	787	4	US-09-721-137-2	Sequence 2, App1
45	35	52.2	787	4	US-09-721-251-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-09-101-272G-62
; Sequence 62, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: residues 43-131 of the ATP domain of uPA
US-09-101-272G-62

Query Match 100.0%; Score 67; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCLP 12
Db 20 ASTDTMGRPCLP 31

RESULT 2
US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guillon, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patricia
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville

STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-12

Query Match 100.0%; Score 67; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
Db 65 ASTDTMGRPCLP 76

RESULT 3
US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. 6686179
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Gulton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 100.0%; Score 67; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
Db 65 ASTDTMGRPCLP 76

RESULT 4
US-08-142-590B-25
Sequence 25, Application US/08142590B
Patent No. 6120765
GENERAL INFORMATION:
APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and
GOETTINCK,
TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,590B
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,318
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match 100.0%; Score 67; DB 3; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 62 ASTDTMGRPCLP 73

RESULT 5
US-09-101-272G-80
; Sequence 80, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 80
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI chimeric protein
US-09-101-272G-80

Query Match 100.0%; Score 67; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 63 ASTDTMGRPCLP 74

RESULT 6
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATP domain of uPA
US-09-101-272G-73

Query Match 100.0%; Score 67; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||

DB 82 ASTDTMGRPCLP 93

RESULT 7
US-09-101-272G-96
; Sequence 96, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 96
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI-CL chimeric protein
US-09-101-272G-96

Query Match 100.0%; Score 67; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 63 ASTDTMGRPCLP 74

RESULT 8
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI-ML chimeric protein
US-09-101-272G-98

Query Match 100.0%; Score 67; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 63 ASTDTMGRPCLP 74

RESULT 9
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No. 5681721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERO J.
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES

APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83

Query Match 100.0%; Score 67; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRCLP 12
DB 16 ASTDTMGRCLP 27

RESULT 10
US-08-720-012-83
Sequence 83, Application US/08720012
Patent No. 5747291
GENERAL INFORMATION:
APPLICANT: STEFFENS, GERD J.
APPLICANT: WENNDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83

Query Match 100.0%; Score 67; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRCLP 12
DB 16 ASTDTMGRCLP 27

RESULT 11
US-08-560-098A-44
Sequence 44, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 100.0%; Score 67; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPLP 12
Db 17 ASTDTMGRPLP 28

RESULT 12

US-08-967-024C-24
Sequence 24, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOWKA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 100.0%; Score 67; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPLP 12
Db 17 ASTDTMGRPLP 28

RESULT 13

US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEFFENS, Gerd Josef

APPLICANT: JANOWKA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 67; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPLP 12
Db 17 ASTDTMGRPLP 28

RESULT 14

US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692
GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewicz, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Faase, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
US-08-087-163-1

Query Match 100.0%; Score 67; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.0007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPCLP 12
Db 62 ASTDTMGRPCLP 73

RESULT 15
US-08-286-748B-18
Sequence 18, Application US/08286748B
Patent No. 5759542
GENERAL INFORMATION:
APPLICANT: Victor Gurewich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
OF DRUGS BY PLATELETS FOR THE TREATMENT OF
TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Faase
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-18

Query Match 100.0%; Score 67; DB 1; Length 411;

Best Local Similarity 100.0%; Pred. No. 0.0007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPCLP 12
Db 62 ASTDTMGRPCLP 73

Search completed: March 18, 2004, 11:21:22
Job time: 1.39168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:19:25 ; Search time 3.57112 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-4
Perfect score: 67
Sequence: 1 ASTDTMGRPCLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	12	13	US-10-076-421-4
2	67	100.0	86	15	US-10-233-675A-1
3	67	100.0	86	15	US-10-233-675A-5
4	67	100.0	86	15	US-10-233-675A-7
5	67	100.0	86	15	US-10-233-675A-22
6	67	100.0	86	15	US-10-233-675A-27
7	67	100.0	87	15	US-10-233-675A-9
8	67	100.0	87	15	US-10-233-675A-10
9	67	100.0	88	9	US-09-880-503-1
10	67	100.0	96	9	US-09-880-503-3
11	67	100.0	135	9	US-09-880-503-4
12	67	100.0	138	9	US-09-984-186-12
13	67	100.0	138	14	US-10-237-667-12
14	67	100.0	138	14	US-10-237-708-12
15	67	100.0	138	14	US-10-237-866-12

16	67	100.0	138	14	US-10-237-871-12	Sequence 12, Appl
17	67	100.0	138	14	US-10-237-624-12	Sequence 12, Appl
18	67	100.0	143	9	US-09-880-503-8	Sequence 8, Appl
19	67	100.0	332	15	US-10-233-675A-20	Sequence 20, Appl
20	67	100.0	332	15	US-10-233-675A-21	Sequence 21, Appl
21	67	100.0	337	14	US-10-106-698-6266	Sequence 6266, Ap
22	67	100.0	337	15	US-10-264-049-2927	Sequence 2927, Ap
23	67	100.0	403	9	US-09-880-503-6	Sequence 6, Appl
24	67	100.0	411	9	US-09-880-503-3	Sequence 3, Appl
25	67	100.0	411	15	US-10-407-821-2	Sequence 2, Appl
26	67	100.0	431	9	US-09-264-468B-1	Sequence 1, Appl
27	67	100.0	431	12	US-10-411-037-34	Sequence 34, Appl
28	67	100.0	431	13	US-10-076-421-2	Sequence 8, Appl
29	67	100.0	431	14	US-10-171-311-184	Sequence 184, App
30	67	100.0	431	14	US-10-193-656-4	Sequence 4, Appl
31	67	100.0	431	14	US-10-301-822-161	Sequence 161, App
32	67	100.0	431	14	US-10-247-671-149	Sequence 149, App
33	67	100.0	431	14	US-10-131-985-21	Sequence 21, Appl
34	67	100.0	431	15	US-10-295-027-414	Sequence 414, App
35	67	100.0	445	15	US-10-295-027-1275	Sequence 1275, Ap
36	67	100.0	445	15	US-10-360-101-266	Sequence 266, App
37	67	100.0	650	15	US-10-401-077-1	Sequence 1, Appl
38	67	100.0	672	15	US-10-233-675A-15	Sequence 15, Appl
39	67	100.0	674	15	US-10-233-675A-14	Sequence 14, Appl
40	67	100.0	687	15	US-10-233-675A-17	Sequence 17, Appl
41	67	100.0	688	15	US-10-233-675A-18	Sequence 18, Appl
42	67	100.0	689	15	US-10-233-675A-13	Sequence 13, Appl
43	51	76.1	86	15	US-10-233-675A-3	Sequence 3, Appl
44	51	76.1	86	15	US-10-233-675A-24	Sequence 24, Appl
45	41	61.2	96	10	US-09-764-872-437	Sequence 437, App

ALIGNMENTS

RESULT 1
US-10-076-421-4
; Sequence 4, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, MANABU
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-4

Query Match 100.0%; Score 67; DB 13; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
DB 1 ASTDTMGRPCLP 12

RESULT 2
US-10-233-675A-1
; Sequence 1, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Neebit, Mark

```
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
; US-10-233-675A-1

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 3
US-10-233-675A-5
; Sequence 5, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
; US-10-233-675A-5

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 4
US-10-233-675A-7
; Sequence 7, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
```

```
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
; US-10-233-675A-7

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 5
US-10-233-675A-22
; Sequence 22, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fragment of human urokinase plasminogen activator
; US-10-233-675A-22

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 6
US-10-233-675A-27
; Sequence 27, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
```


SOFTWARE: Patentin version 3.1
SEQ ID NO 27
LENGTH: 86
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fragment of human urokinase plasminogen activator
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (43)-(43)
OTHER INFORMATION: Polymorphism - Xaa = Asn or Asp
NAME/KEY: MISC FEATURE
LOCATION: (74)-(74)
OTHER INFORMATION: Polymorphism - Xaa = Pro or Leu
US-10-233-675A-27

Query Match 100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 15 ASTDTMGRPCLP 26

RESULT 7
US-10-233-675A-9
Sequence 9, Application US/10233675A
Publication No. US20030228298A1
GENERAL INFORMATION:
APPLICANT: Neabtic, Mark
APPLICANT: Fong, Timothy
TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
FILE REFERENCE: ST01027
CURRENT APPLICATION NUMBER: US/10/233,675A
CURRENT FILING DATE: 2002-09-04
PRIOR APPLICATION NUMBER: 60/316,300
PRIOR FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin version 3.1
SEQ ID NO 9
LENGTH: 87
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: human derived fusion protein
US-10-233-675A-9

Query Match 100.0%; Score 67; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 16 ASTDTMGRPCLP 27

RESULT 8
US-10-233-675A-10
Sequence 10, Application US/10233675A
Publication No. US20030228298A1
GENERAL INFORMATION:
APPLICANT: Neabtic, Mark
APPLICANT: Fong, Timothy
TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
FILE REFERENCE: ST01027
CURRENT APPLICATION NUMBER: US/10/233,675A
CURRENT FILING DATE: 2002-09-04

PRIOR APPLICATION NUMBER: 60/316,300
PRIOR FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin version 3.1
SEQ ID NO 10
LENGTH: 87
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: human derived fusion protein
US-10-233-675A-10

Query Match 100.0%; Score 67; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 16 ASTDTMGRPCLP 27

RESULT 9
US-09-880-503-1
Sequence 1, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 88
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-1

Query Match 100.0%; Score 67; DB 9; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 15 ASTDTMGRPCLP 26

RESULT 10
US-09-880-503-9
Sequence 9, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 96
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-9

Query Match 100.0%; Score 67; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
Db 15 ASTDTMGRPCLP 26

RESULT 11
US-09-880-503-4
Sequence 4, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
FILE REFERENCE: 9596-331
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 100.0%; Score 67; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
Db 62 ASTDTMGRPCLP 73

RESULT 12
US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. US20020151011A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689

FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 100.0%; Score 67; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
Db 65 ASTDTMGRPCLP 76

RESULT 13
US-10-237-667-12
Sequence 12, Application US/10237667
Publication No. US20030022308A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12

Query Match 100.0%; Score 67; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 65 ASTDTMGRPCLP 76

RESULT 14
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph. D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 100.0%; Score 67; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 65 ASTDTMGRPCLP 76

RESULT 15
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph. D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 100.0%; Score 67; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||
Db 65 ASTDTMGRPCLP 76

Search completed: March 18, 2004, 11:36:07
Job time : 3.82112 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 1.23414 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-4

Perfect score: 67

Sequence: 1 ASTDTWGRPLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :
1: PIR 78: *
2: p1r1: *
3: p1r2: *
4: p1r3: *
5: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	431	1 UKHU	u-plasminogen acti
2	52	77.6	433	1 UKBAY	u-plasminogen acti
3	51	76.1	432	1 S18932	u-plasminogen acti
4	51	76.1	443	1 UKMS	u-plasminogen acti
5	51	76.1	442	1 UKPG	u-plasminogen acti
6	40	59.7	224	2 S20463	siderophore biosyn
7	40	59.7	500	1 JC4153	cytochrome P450 2D
8	39	58.2	27	2 S29211	probable phosphol
9	39	58.2	288	2 T38696	probable peroxisom
10	39	58.2	419	2 S42389	T48 protein - fru
11	39	58.2	433	1 JN0560	u-plasminogen acti
12	39	58.2	543	2 AF1927	hypothetical prote
13	39	58.2	618	2 A35827	thrombin (EC 3.4.2
14	38	56.7	125	2 T49092	SB-like protein -
15	38	56.7	167	2 H72579	hypothetical prote
16	38	56.7	437	2 H86683	prophage p11 prote
17	38	56.7	589	2 I59297	parathyroid hormon
18	38	56.7	591	2 S44203	parathyroid hormon
19	38	56.7	591	2 I54195	parathyroid hormon
20	38	56.7	593	2 A49191	parathyroid hormon
21	37	55.2	153	2 B46353	ORF2 protein - COC
22	37	55.2	409	1 B0A655	virB10 protein - A
23	37	55.2	773	2 A83888	hypothetical prote
24	37	55.2	804	2 T44606	transducer protein
25	37	55.2	810	2 T46810	halobacterial tran
26	37	55.2	810	2 P84327	Htr5 transducer [I
27	37	55.2	1299	1 WMBEHS	membrane antigen p
28	37	55.2	1348	2 S27812	probable epidermal
29	37	55.2	1348	2 A43917	probable epidermal

30	36	53.7	231	2 H70968	hypothetical prote
31	36	53.7	247	2 T37820	hypothetical prote
32	36	53.7	322	2 E89057	protein K09H1.7 [
33	36	53.7	335	2 T31712	hypothetical prote
34	36	53.7	349	2 T15082	hypothetical prote
35	36	53.7	353	2 S05297	regulatory protein
36	36	53.7	374	2 F86243	ZIP4, probable zin
37	36	53.7	396	2 A82532	translation elonga
38	36	53.7	438	2 I50517	retinoid X recepto
39	36	53.7	529	2 B24059	t complex polypept
40	36	53.7	533	2 T45197	probable serine pr
41	36	53.7	556	2 S13163	t-complex-type mol
42	36	53.7	556	2 S10486	t-complex-type mol
43	36	53.7	556	2 JC1443	t complex polypept
44	36	53.7	556	2 JQ0866	T-complex protein
45	36	53.7	812	2 A46417	NPI1 protein - yea

ALIGNMENTS

RESULT 1
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N:Contents: urokinase-type plasminogen activator chain A; urokinase-type plasminogen acti
in form
C:Species: Homo sapiens (man)
C>Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C/Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37563; A37564; A356
R/Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasl, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A>Title: The human urokinase-plasminogen activator gene and its promoter.
A/Reference number: A00931; M01D:85215647; P01D:2587867
A/Accession: A00931
A/Molecule type: DNA
A/Residues: 1-431 <RIC>
A/Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:G134524
A/Note: The authors translated the codon ATG for residue 214 as Ile
R/Magamine, Y.; Pearson, D.; Gratian, M. 1985
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A>Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine u
R/Nagai, M.; Hiratsugu, R.; Kaneda, T.; Hayasuke, N.; Nishida, M.; Suyama, T
Gene 36, 183-188, 1985
A>Title: Molecular cloning of cDNA coding for human preprourokinase.
A/Reference number: J70102; M01D:86056394; P01D:3933505
A/Accession: I52209
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 145-161 <NAG1>
A/Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R/Nagai, M.; Hiratsugu, R.; Kaneda, T.; Hayasuke, N.; Nishida, M.; Suyama, T
Gene 36, 183-188, 1985
A>Title: Molecular cloning of cDNA coding for human preprourokinase.
A/Reference number: J70102; M01D:86056394; P01D:3933505
A/Accession: I52209
A/Molecule type: mRNA
A/Residues: 1-213, 'I', 215-431 <NAG2>
A/Cross-references: GB:X03026; NID:G340155; PIDN:AA971138.1; PID:G340158; GB:D00244; NID
R/Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasl, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A>Title: Identification and primary sequence of an unspliced human urokinase poly(A) + RN
A/Reference number: A37561; M01D:84272706; P01D:6589620
A/Accession: A37561
A/Molecule type: mRNA
A/Residues: 66-431 <VER>
A/Cross-references: GB:D00244; NID:G220138
R/Jacobs, P.; Cravador, A.; Lortie, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elee
DNA 4, 133-146, 1985
A>Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prepr
A/Reference number: I38102; M01D:85203359; P01D:3888571
A/Accession: I38102
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A/Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

R.Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
 A:Title: Characterization of single chain urokinase-type plasminogen activator with a no
 A:Reference number: S65783; MUID:96186279; PMID:8652631
 A:Accession: S65783
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 21-140, 'V', 142-213, 'T', 215-431 <YOS>
 A:Cross-references: EMBL:D1143; NID:9311467; PIDN:BA01919.1; PID:g119928
 R:Guzlier, W.A.; Steffens, G.U.; Oetting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A:Title: The primary structure of high molecular mass urokinase from human urine.
 A:Reference number: A37562; MUID:83055084; PMID:6754569
 A:Accession: A37562
 A:Molecule type: protein
 A:Residues: 21-177 <GUN>
 R:Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Jergler, W.; Studer, R.O.
 Eur. J. Biochem. 125, 251-257, 1982
 A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and preli
 A:Reference number: A37563; MUID:83003608; PMID:6749491
 A:Accession: A37563
 A:Molecule type: protein
 A:Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>
 R:Steffens, G.U.; Gunzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A:Title: The complete amino acid sequence of low molecular mass urokinase from human uri
 A:Reference number: A37564; MUID:83055099; PMID:6754572
 A:Accession: A37564
 A:Molecule type: protein
 A:Residues: 158-410 <STE>
 R:Kertzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A:Title: Carboxylate composition and presence of a fucose-protein linkage in recombinat
 A:Reference number: A35689; MUID:90365737; PMID:2393398
 A:Accession: A35689
 A:Molecule type: protein
 A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
 A:Note: Identification of a fucose and attempt to determine its attachment site
 R:Rabhand, S.A.; Desjardins, J.; Bell, A.W.; Bannville, D.; Mazur, A.; Henkin, J.; Goltz
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
 A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
 A:Reference number: A36697; MUID:91097529; PMID:2125213
 A:Accession: A36697
 A:Molecule type: protein
 A:Residues: 21-34 <RAB>
 R:Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
 Submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A51255; PDB:1KDU
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Li, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
 A:Reference number: A44375; MUID:93003110; PMID:1327118
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R:Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
 submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A66822; PDB:1URK
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A66058; PDB:1LMM
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175, 179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C:Genetics:
 A:Gene: GDB:PLAU
 A:Cross-references: GDB:119497; OMIM:191840
 A:Map position: 10q24-10q24
 A:Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteinat
 F:1-20/Domains: signal sequence #status predicted <SIG>
 F:1-431/Product: urokinase-type plasminogen activator, single chain form #status predict
 F:1-177/Product: urokinase-type plasminogen activator chain A #status experimental <MPA>
 F:31-62/Domains: EGF homology <EGF>
 F:70-151/Domains: kringle homology <KR>
 F:156-177/Product: urokinase-type plasminogen activator chain B #status experimental <MP
 F:179-431/Product: urokinase-type plasminogen activator chain A #status experimental <MP
 F:179-431/Domains: trypsin homology <TRY>
 F:31-39,33-51,53-62,70-151,91-133,132-146,168-299,209-225,217-288,313-382,345-361,372-400
 F:38/Binding site: carboxylate (Thr) (covalent) #status predicted
 F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
 F:224,275,376/Active site: His, Asp, Ser #status experimental
 F:322/Binding site: carboxylate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 67; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.00019;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPLP 12
 |||||
 Db 82 ASTDTMGRPLP 93

RESULT 2

UKBAY

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999

C:Accession: S14687; S08651.

R:Yu, Y.P.T.; Wang, T.W.; Clowes, A.W.

Nucleic Acids Res. 18, 3411, 1990

A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen

A:Reference number: S14687; MUID:90287734; PMID:2113276

A:Accession: S14687

A:Molecule type: mRNA

A:Residues: 1-433 <AY>

A:Cross-references: EMBL:X51935; NID:938130; PIDN:CA36200.1; PID:938131

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F:1-20/Domains: signal sequence #status predicted <SIG>

F:21-176/Product: plasminogen activator chain A #status predicted <ACH>

F:30-61/Domains: EGF homology <EGF>

F:69-150/Domains: kringle homology <KR>

F:178-433/Product: plasminogen activator chain B #status predicted <BC>

F:178-431/Domains: trypsin homology <TRY>

F:167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted

F:223,274,378/Active site: His, Asp, Ser #status predicted

F:324/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 77.6%; Score 52; DB 1; Length 433;
 Best Local Similarity 90.9%; Pred. No. 0.1;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPLP 11
 |||||
 Db 81 ASTDTMGRSCL 91

RESULT 3

S18932

u-plasminogen activator (EC 3.4.21.73) precursor - rat

N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999

C:Accession: S24604; I60186; I53472; S18932

R:Rabhand, S.A.

submitted to the EMBL Data Library, April 1992

A:Reference number: S24604

A:Accession: S24604

A:Molecule type: mRNA

A:Residues: 1-15, 'H', 17-23, 'G', 25-33, 'N', 33-432 <RAB>

A:Cross-references: EMBL:X65651; NID:957456; PIDN:CAA4601.1; PID:957457

A, Experimental source: tissue kidney
 R: Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
 Cancer Res. 52, 2489-2496, 1992
 A, Title: Transcriptional and posttranscriptional activation of urokinase plasminogen acti
 A, Reference number: 160186; MUID:92233409; PMID:1568219
 A, Accession: 160186
 A, Status: preliminary; translated from GB/EMBL/DBJ
 A, Molecule type: mRNA
 A, Residues: 1-432 <RES>
 A, Cross-references: EMBL:X63434; NID:957465; PIDN:CAA45028.1; PID:957466
 A, Experimental source: strain Fischer 344; tissue mammary
 R: Rago, F.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
 FEBS Lett. 306, 193-198, 1992
 A, Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
 A, Reference number: 153472; MUID:92339549; PMID:1321734
 A, Accession: 153472
 A, Status: preliminary; translated from GB/EMBL/DBJ
 A, Molecule type: DNA
 A, Residues: 31-62 <RES2>
 A, Cross-references: EMBL:X66907; NID:9396200; PIDN:CAA47356.1; PID:9398279
 C: Genes: uPA
 C: Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C: Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F: 1-19/Domain: signal sequence #status predicted <SIG>
 F: 20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F: 31-62/Domain: EGF homology <EGF>
 F: 70-151/Domain: kringle homology <KR>
 F: 179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F: 179-420/Domain: trypsin homology <TRY>
 F: 168-300, 210-226, 218-289, 314-363, 373-401/Disulfide bonds: #status predicted
 F: 225, 276, 377/Active site: His, Asp, Ser #status predicted

Query Match 76.1%; Score 51; DB 1; Length 432;
 Best Local Similarity 81.8%; Pred. No. 0.16;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 11
 Db 82 ANNTDTCGRPCL 92

RESULT 4
 UKPG
 u-plasminogen activator (EC 3.4.21.73) precursor - mouse
 C: Species: Mus musculus (house mouse)
 C: Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
 C: Accession: A29420; A24615
 R: Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 Biochemistry 26, 8270-8279, 1987
 A, Title: The murine urokinase-type plasminogen activator gene.
 A, Reference number: A29420; MUID:88163489; PMID:2831940
 A, Accession: A29420
 A, Molecule type: DNA
 A, Residues: 1-433 <DEG>
 A, Cross-references: GB:M17922; NID:9202296; PIDN:AAA40539.1; PID:9202297
 R: Bellin, D.; Vassalli, J.D.; Combarnet, C.; Goddard, F.; Nagnan, Y.; Reich, E.; Kocher,
 Eur. J. Biochem. 148, 225-232, 1995
 A, Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
 A, Reference number: A24615; MUID:85179474; PMID:2985383
 A, Accession: A24615
 A, Molecule type: DNA
 A, Residues: 1-433 <BNL>
 A, Cross-references: GB:X02389; NID:955127; PIDN:CAA26331.1; PID:955128
 C: Genes: uPA
 A, Intons: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C: Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C: Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F: 1-20/Domain: signal sequence #status predicted <SIG>
 F: 21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F: 32-63/Domain: EGF homology <EGF>
 F: 71-152/Domain: kringle homology <KR>
 F: 180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F: 180-421/Domain: trypsin homology <TRY>
 F: 159-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
 F: 226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 76.1%; Score 51; DB 1; Length 433;
 Best Local Similarity 81.8%; Pred. No. 0.16;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 11
 Db 83 ANNTDTCGRPCL 93

RESULT 5
 UKPG
 u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N: Alternate names: uPA
 C: Species: Sus scrofa domestica (domestic pig)
 C: Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C: Accession: A00932
 R: Nagnan, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A, Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A, Reference number: A00932; MUID:85087954; PMID:6096832
 A, Accession: A00932
 A, Molecule type: DNA
 A, Residues: 1-240; 'H', 242-442 <NAG1>
 A, Experimental source: kidney cell line LLC-PK1
 R: Nagnan, Y.
 A, Submitted to the Protein Sequence Database, December 1986
 A, Reference number: A37566
 A, Contents: annotation; correction to residue 241
 C: Genes: uPA
 A, Intons: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C: Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C: Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F: 1-20/Domain: signal sequence #status predicted <SIG>
 F: 21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F: 33-64/Domain: EGF homology <EGF>
 F: 72-153/Domain: kringle homology <KR>
 F: 180-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F: 190-430/Domain: trypsin homology <TRY>
 F: 152/Binding site: carboxylate (Asn) (covalent) #status predicted
 F: 179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
 F: 235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 76.1%; Score 51; DB 1; Length 442;
 Best Local Similarity 75.0%; Pred. No. 0.16;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 12
 Db 84 ANNTDTCGRPCL 95

RESULT 6
 S20463
 siderophore biosynthesis regulatory protein sfp(0) - Bacillus subtilis
 N: Alternate names: surfactin producing protein
 C: Species: Bacillus subtilis
 C: Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Jun-2000
 C: Accession: S20463; A36931; F69705; S26594
 R: Nakano, M.M.; Corbell, N.; Besson, J.; Zuber, P.
 Mol. Gen. Genet. 232, 313-321, 1992
 A, Title: Isolation and characterization of sfp: a gene that functions in the production
 A, Reference number: S20463; MUID:92212296; PMID:1557038
 A, Accession: S20463
 A, Molecule type: DNA
 A, Residues: 1-224 <NAX>
 A, Cross-references: EMBL:X63158; NID:940138; PIDN:CAA44858.1; PID:940139
 A, Experimental source: strain OKB105
 R: Grossman, T.H.; Tuckman, M.; Ellstedt, S.; Osburne, M.S.
 J. Bacteriol. 175, 6203-6211, 1993

C:/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Nov-1999
C:/Accession: S42989
R:/Strut, D.I.; White, R.A.H.
submitted to the EMBL Data Library, March 1994
A:/Description: Characterisation of T48, a target of homeotic gene regulation in *Drosophila*
A:/Reference number: S42989
A:/Accession: S42989
A:/Status: preliminary
A:/Molecule type: mRNA
A:/Residues: 1-419 <STR>
A:/Cross-references: EMBL:X78113; NID:G459593; PID:G459594
C:/Genetics:
A:/Gene: FlyBase:T48
A:/Cross-references: FlyBase:FBgn0004359
C:/Superfamily: fruit fly T48 protein

Query Match 58.2%; Score 39; DB 2; Length 419;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Ox 1 ASTDTMGRCPL 12
||| |||
Db 135 AHTEFMGRVCP 146

RESULT 11
UN0560 u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N:/Alternate names: uPA
C:/Species: Bos primigenius taurus (cattle)
C:/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:/Accession: UN0560
R:/Kraetschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A:/Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and inducibility
A:/Reference number: UN0560; MUID:93216119; PMID:8385052
A:/Accession: UN0560
A:/Molecule type: mRNA
A:/Residues: 1-433 <KKA>
A:/Cross-references: GB:L03546; NID:G163800; PID:NAA51419.1; PID:G163801
C:/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; tryk
C:/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:/1-179/Product: signal sequence #status predicted <SIG>
F:/21-179/Product: plasminogen activator chain A #status predicted <MA1>
F:/33-64/Domains: EGF homology <EGF>
F:/72-153/Domains: kringle homology <KR>
F:/181-433/Product: plasminogen activator chain B #status predicted <MA2>
F:/181-433/Domains: trypsin homology <TRY>
F:/170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F:/226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 58.2%; Score 39; DB 1; Length 433;
Best Local Similarity 63.6%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Ox 1 ASTDTMGRCPL 11
||| |||
Db 84 ANRDLSGRPCL 94

RESULT 12
AF1927 hypotheical protein alio969 [imported] - Noetic sp. (strain PCC 7120)
C:/Species: Noetic sp. PCC 7120
A:/Note: Noetic sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:/Accession: AF1927
R:/Keneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchihara,
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
C:/Reference number: AB1807; MUID:21595285; PMID:11759840

```

A:Accession: AF1927
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <KDR>
A:Cross-references: GB:BA000019; PIDN:BAB72926.1; PID:gl7130315; GSPDB:GN00179
C:Genetics:
A:Experimental source: strain PCC 7120
A:Gene: all0969

Query Match      58.2% Score 39; DB 2; Length 543;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 ASTDTMRPCLP 12
Db      230 ANTDTQGPFPF 241

RESULT 13
A35827
Chrombin (EC 3.4.21.5) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 14-Dec-1990 #sequence revision 14-Dec-1990 #text_change 03-May-2002
C:Accession: A35827; A42696; S12081
R:Degen, S.J.F.; Scheffer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J.A.
DNA Cell Biol. 9, 487-499, 1990
A:Title: Characterization of the cDNA coding for mouse prothrombin and localization of th
A:Reference number: A35827; MUID:91025551; PMID:2222810
A:Accession: A35827
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-618 <DBS>
A:Cross-references: GB:X52308; NID:953813; PIDN:CA36548.1; PID:953814
A:Experimental source: strain C57BL/6
A:Note: the data were obtained from females resulting from the cross of M. domesticus and
R:Banfield, D.K.; MacCullivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequ
A:Reference number: A42696; MUID:92212913; PMID:1557883
A:Accession: A42696
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 384-618, 'E', <BAN>
A:Cross-references: GB:M61394
C:Superfamily: chrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydr
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-43/Domain: propeptide #status predicted <PRO>
F:28-88/Domain: Gla domain homology <Gla>
F:44-618/Product: prothrombin B #status predicted <MNT>
F:109-187/Domain: kringle homology <KR1>
F:215-293/Domain: kringle homology <KR2>
F:361-610/Domain: trypsin homology <TRY>
E:50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F:61-66,91-104,109-181,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,54
F:403,459/Active site: His, Asp, Ser #status predicted

Query Match      58.2% Score 39; DB 2; Length 618;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      5 TMGRPCLP 12
Db      231 TLGSPCLP 238

RESULT 14
T49092
5B-like protein - Arabidopsis thaliana
N:Alternate names: protein PAP15.240
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
A:Accession: T49092

```

R;Alcaraz, J.P.; Clabault, G.; Cottet, A.; Mache, R.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25015
A;Accession: T49092
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-125 <ALC>
A;Cross-references: EMBL:AL049711; GSPDB:GN00061; ATSP:F4F15.240
A;Experimental source: cultivar Columbia; BAC clone F4F15
C;Genetics:
A;Gene: ATSP:F4F15.240
A;Map position: 3

Query Match 56.7%; Score 38; DB 2; Length 125;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 5 TMGRPCL 11
|:|||||
Db 84 TLGRPCL 90

RESULT 15

H72579
hypothetical protein APE1920 - Aeropyrum pernix (strain KI)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: H72579
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999.
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72579
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-167 <KAW>
A;Cross-references: DDBJ:AP000062; NID:G5105244; PIDN:BAA80925.1; PID:dl044711; PID:G510
A;Experimental source: strain KI
C;Genetics:
A;Gene: APE1920

Query Match 56.7%; Score 38; DB 2; Length 167;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TMGRPCLP 12
|:|||||
Db 64 TFGRPCLP 71

Search completed: March 18, 2004, 11:20:18
Job time : 2.23414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2004, 10:54:45 ; Search time 0.73523 Seconds

(without alignments)
849.859 Million cell updates/sec

Title: US-10-076-421-4

Perfect score: 67

Sequence: 1 ASTDTNGRPLCP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	431	1	UROC_HUMAN
2	52	77.6	433	1	UROC_PAPCY
3	51	76.1	432	1	UROC_RAT
4	51	76.1	433	1	UROC_MOUSE
5	51	76.1	442	1	UROC_PIG
6	40	59.7	224	1	SFP_BACSU
7	40	59.7	500	1	CPDG_CAVPO
8	39	58.2	433	1	UROC_BOVIN
9	39	58.2	602	1	YHS4_CABEL
10	38	56.7	618	1	THRB_MOUSE
11	38	56.7	585	1	PTRR_PIG
12	38	56.7	591	1	PTRR_MOUSE
13	38	56.7	591	1	PTRR_RAT
14	38	56.7	593	1	PTRR_HUMAN
15	37	55.2	205	1	PSB3_TRYBB
16	37	55.2	376	1	VIBX_AGRTO
17	37	55.2	377	1	VIBX_AGRTO
18	37	55.2	810	1	HTR4_HALNI
19	37	55.2	810	1	HTR4_HALSA
20	37	55.2	1299	1	VP75_HSVSA
21	37	55.2	1693	1	SAS_DROME
22	36	53.7	230	1	EIA_ADECC
23	36	53.7	247	1	YE95_SCHPO
24	36	53.7	353	1	NIFA_RHLIT
25	36	53.7	374	1	ZIP4_ARATH
26	36	53.7	395	1	EFTU_XYLFA
27	36	53.7	395	1	EFTU_XYLFA
28	36	53.7	457	1	TMS5_HUMAN
29	36	53.7	556	1	TCP2_MOUSE
30	36	53.7	556	1	TCPA_CRIGR
31	36	53.7	556	1	TCPA_HUMAN
32	36	53.7	556	1	TCPA_RAT
33	36	53.7	559	1	TCPA_PALPA

34	36	53.7	812	1	IF38_YEAST
35	36	53.7	864	1	AGLU_MUCJA
36	36	53.7	954	1	YB79_YEAST
37	36	53.7	3164	1	TEGU_HSV11
38	35	52.2	126	1	WN72_EPRST
39	35	52.2	224	1	SFP_BACSU
40	35	52.2	477	1	NIFE_RHOCA
41	35	52.2	530	1	CP14_CHICK
42	35	52.2	630	1	KCD2_HUMAN
43	35	52.2	690	1	AFT1_YEAST
44	35	52.2	696	1	SYM_XANCP
45	35	52.2	726	1	CNG2_ARATH

ALIGNMENTS

RESULT 1
UROC_HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q16618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human Urokinase-plasminogen activator gene and its promoter";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Roy M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";
RN Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=6056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RT "Molecular cloning of cDNA coding for human prepro-urokinase";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Crayador A., Lorlau R., Brockly F., Colau B., Chuchana P.,
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA";
RL DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi O.,
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=23288257; PubMed=12477932;
RA Stranberg R.L., Reinhold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RT Altschuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Caninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman A.C., Shevchenko Y., Bouffard G.G.,
 RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
 RA Blakesley R.W., Touchman J.W., Schmitt J., Myers R.M.,
 RA Rodriguez A.C., Grimwood J., Skalska U., Smallus D.E.,
 RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
 RL [7]
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blaesi F.,
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA".
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RL [8]
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.,
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain".
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RL [9]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749991;
 RA Schaller O., Nick H., Rickli E.E., Gillissen D., Lergier W.,
 RA Studer R.O.,
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains".
 RL Eur. J. Biochem. 125:251-257(1982).
 RL [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.,
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine".
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RL [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.,
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator".
 RL Structure 3:681-691(1995).
 RL [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Speil S., Jacob U., Arroyo de Frada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.,
 RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase".
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RL [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Boguski M.J., Bamberger M., Smith R.A.G., Dobson C.M.,
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR".
 RL Nature 337:579-582(1989).
 RL [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.,

RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase".
 RL Biochemistry 31:9562-9571(1992).
 RL [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=9419701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.,
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator".
 RL J. Mol. Biol. 235:1548-1559(1994).
 RL [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Yamaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hamada K.,
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure".
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RL [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Comme B., Berczy M., Belin D.,
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene".
 RL Thromb. Haemost. 77:434-435(1997).
 RL [18]
 RP ERRATUM.
 RA Comme B., Berczy M., Belin D.,
 RL Thromb. Haemost. 78:973-973(1997).
 RL [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalzfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graef H., Magdolen V.,
 RT "Mutational analysis of the gene encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer".
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: Potent plasminogen activator and is clinically used for
 CC therapy of thrombolytic disorders.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain.
 CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in pulmonary embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X02419; CAA26268.1; -;
 DR EMBL; M15476; AAA61253.1; -;
 DR EMBL; D00244; BAA00175.1; -;
 DR EMBL; D11143; BAA01919.1; -;
 DR EMBL; X02760; CAA26535.1; -;
 DR EMBL; AF377330; AAK53822.1; -;
 DR EMBL; BC013575; AAH13575.1; -;
 DR EMBL; X03226; AAC97138.1; -;
 DR EMBL; K02286; AAA61252.1; -;
 DR EMBL; A21571; CAA01559.1; -;
 DR EMBL; A18397; CAA01390.1; -;
 DR PIR; A00931; UKHD.
 DR PDB; 1KDU; 31-OCT-93.

Query Match 100.0%; Score 67; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASTDTMGRPCLP 12
 DB 82 ASTDTMGRPCLP 93

RESULT 2
 UROK_PAPCY STANDARD; PRT; 433 AA.
 AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Papio cynocephalus (Yellow baboon).
 OS Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecidae; Papio.
 OX NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA Xu Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RT plasminogen activator";
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1-chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.jsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: X51935; CA36200.1; -.
 DR PIR: S14687; UKBAY.
 DR HSSP: P00749; ILMW.
 DR MEROPS: S01.231; -.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR006209; EGF_Like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR008293; Pept_S1A_uPA.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00051; Kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PIRSF: PIRSF001144; UR_plasm_act; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_Spec; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; FALSE_NEG.

DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL
 FT CHAIN 1
 FT CHAIN 20
 FT CHAIN 21 433
 FT CHAIN 21 176
 FT CHAIN 155 176
 FT CHAIN 178 433
 FT CHAIN 26 62
 FT DOMAIN 69 150
 FT DOMAIN 151 177
 FT DOMAIN 178 433
 FT DISULFID 30 38
 FT DISULFID 32 50
 FT DISULFID 52 61
 FT DISULFID 167 298
 FT DISULFID 208 224
 FT DISULFID 216 287
 FT DISULFID 315 384
 FT DISULFID 347 363
 FT DISULFID 374 402
 FT ACT_SITE 223 223
 FT ACT_SITE 274 274
 FT ACT_SITE 378 378
 FT CARBOHYD 324 324
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDC8792 CRC64; (BY SIMILARITY).

Query Match 77.6%; Score 52; DB 1; Length 433;
 Best Local Similarity 90.9%; Pred. No. 0.038; 1; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ASTDTMGRPCLP 11
 DB 81 ASTDTMGRPCLP 91

RESULT 3
 UROK_RAT STANDARD; PRT; 432 AA.
 AC P29598;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer 344;
 RX MEDLINE=92233409; PubMed=1568219;
 RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
 RA Kefford R.F.;
 RT "Transcriptional and posttranscriptional activation of urokinase
 RT plasminogen activator gene expression in metastatic tumor cells";
 RT Cancer Res. 52:2489-2496(1992).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a

CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X63434; CAA45028.1; -;
 DR EMBL; X65651; CAA46601.1; -;
 DR PIR; S24604; S18932.
 DR HSSP; P00749; 1KDU.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR01254; Peptidase_S1.
 DR InterPro; IPR01314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM0020; TRYP_Spec; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 432
 FT CHAIN 20 177
 FT CHAIN 156 177
 FT CHAIN 179 432
 FT DOMAIN 27 63
 FT DOMAIN 70 151
 FT DOMAIN 152 178
 FT DOMAIN 179 432
 FT DISULFID 31 39
 FT DISULFID 33 51
 FT DISULFID 53 62
 FT DISULFID 168 300
 FT DISULFID 210 226
 FT DISULFID 218 289
 FT DISULFID 314 383
 FT DISULFID 346 362
 FT DISULFID 373 401
 FT ACT_SITE 225 225
 FT ACT_SITE 276 276
 FT ACT_SITE 377 377
 FT CONFLICT 16 16
 FT CONFLICT 24 24
 FT CONFLICT 332 332
 FT SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;
 Query Match 76.1%; Score 51; DB 1; Length 432;
 Best Local Similarity 81.8%; Fred. No. 0.058;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ASTDTMGRPCL 11
 Db 82 ANTDTKGRPCL 92
 RESULT 4
 ID UROK_MOUSE STANDARD; PRT; 433 AA.
 AC P06869;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (BC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxId=10090;
 RX MEDLINE=85179474; PubMed=2985383;
 RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
 RA Reich E., Koehler H.P., Duvoisin R.M.;
 RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
 RT mouse urokinase-type plasminogen activator.";
 RL Eur. J. Biochem. 148:225-232(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88163489; PubMed=2831940;
 RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
 RT "The murine urokinase-type plasminogen activator gene.";
 RL Biochemistry 26:8270-8279(1987).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02389; CAA26231.1; -;
 DR EMBL; M17922; AAA40539.1; -;
 DR PIR; A29420; UROKS.
 DR HSSP; P00749; 1KDU.
 DR MEROPS; S01.231; -;
 DR MGD; MGI.97611; Plau.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR01254; Pept_S1A_uPA.
 DR InterPro; IPR008293; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.

RESULT 6
SFP_BACSU STANDARD; PRT; 224 AA.
ID_SFP_BACSU
AC P39135;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 4'-phosphopantetheinyl transferase sfp (EC 2.7.8.-) (Surfactin synthetase-activating enzyme).
GN SFP OR LPA-8 OR BSU03570.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=1423;
[1]
RN NCBI
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14.
RC STRAIN=168;
RX MEDLINE=92212296; PubMed=1557038;
RA Nakano M.M., Corbell N., Besson J., Zuber P.;
RT "Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in Bacillus subtilis.";
RL Mol. Genet. 232:313-321 (1992).
[2]
RN NCBI
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH624;
RA Grandi G.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
[3]
RN NCBI
RP SEQUENCE FROM N.A.
RX MEDLINE=94012481; PubMed=8407792;
RA Grossman T.H., Tuckman M., Ellestad S., Osburne M.S.;
RT "Isolation and characterization of Bacillus subtilis genes involved in siderophore biosynthesis: relationship between B. subtilis sfp and Escherichia coli end genes.";
RL J. Bacteriol. 175:6203-6211 (1993).
[4]
RN NCBI
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=93360813; PubMed=8355609;
RA Cosmina P., Rodriguez F., de Ferra F., Grandi G., Perego M., Venema G., van Sinderen D.;
RT "Sequence and analysis of the genetic locus responsible for surfactin synthesis in Bacillus subtilis.";
RL Mol. Microbiol. 8:821-831 (1993).
[5]
RN NCBI
RP SEQUENCE FROM N.A.
RC STRAIN=YB8;
RX MEDLINE=96214657; PubMed=8639027;
RA Tauge K., Ano T., Shoda M.;
RT "Isolation of a gene essential for biosynthesis of the lipopeptide antibiotics plipastatin B1 and surfactin in Bacillus subtilis YB8.";
RL Arch. Microbiol. 165:243-251 (1996).
[6]
RN NCBI
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriss R., Boursier L., Brans A., Braun W., Brigneau S.C., Bron S., Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A., Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blandhard M., Klein C., Kobayashi Y., Koester P., Koningsheim G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maul C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassarotti A., Viari A., Wambuit R., Wedler E., Weitzenger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.";
RL Nature 390:249-256 (1997).
[7]
RN NCBI
RP FUNCTION
RX MEDLINE=97144264; PubMed=8939709;
RA Limalot R.H., Gehring A.M., Flugel R.S., Zuber P., Laccelle M., Marahel M.A., Reid R., Khosla C., Walsh C.T.;
RT "A new enzyme superfamily - the phosphopantetheinyl transferases.";
RL Chem. Biol. 3:923-936 (1996).
[8]
RN NCBI
RP CHARACTERIZATION, AND MUTAGENESIS OF GLX-105; ASF-107; TRP-147;
RX GLU-151 AND LYS-155.
RX MEDLINE=98145223; PubMed=9484229;
RA Quadri L.E.N., Weinreb P.H., Lei M., Nakano M.M., Zuber P., Walsh C.T.;
RT "Characterization of sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases.";
RL Biochemistry 37:1585-1595 (1998).
[9]
RN NCBI
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=20050577; PubMed=10581256;
RA Reuter K., Wolf M.R., Marahel M.A., Fichter R.;
RT "Crystal structure of the surfactin synthetase-activating enzyme sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily.";
RL EMBO J. 18:6823-6831 (1999).
[10]
RN NCBI
RP -1- FUNCTION: Activates the seven peptidyl carrier protein (PCP) domains of surfactin synthetase SFP1/2/3 by transferring the 4'-phosphopantetheinyl moiety of coenzyme A (CoA) to a serine residue. Required for cells of B. subtilis to become producers of the lipopeptide antibiotics surfactin and plipastatin B1.
CC -1- CATALYTIC ACTIVITY: CoA + apo-(peptidyl-carrier protein) = adenosine 3',5'-bisphosphate + holo-(peptidyl-carrier protein).
CC -1- COFACTOR: Magnesium.
CC -1- SUBUNIT: Monomer in solution.
CC -1- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY.
CC GSP/SFP/HEIT/ACPT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; X63158; CAA44858.1; -;
CC EMBL; X65610; CAA46561.1; -;
CC EMBL; L17438; AAC36829.1; -;
CC EMBL; X70356; -; NOT ANNOTATED_CDS.
CC EMBL; D50562; BAA09125.1; -;
CC EMBL; Z99105; CAB12151.1; -;
CC PIR; S20463; S20463.
CC PIR; IQR0; 10-DEC-99.
CC Subtilisin, Bg10176; sfp.
CC Interpro: IPR008278; 4-PPT transf.
CC Interpro: IPR004568; Pantethn_tm.
CC Pfam; PF01648; ACP5; 1.
CC TIGRFAMs; TIGR00556; pantethn_tm; 1.
CC Transferase; Antibiotic biosynthesis; Magnesium; 3D-structure;
KW

Complete proteome.

FT	METAL	107	107	MAGNESIUM.
FT	METAL	109	109	MAGNESIUM.
FT	METAL	151	151	MAGNESIUM.
FT	BINDING	158	189	PCP DOMAIN OF SUBSTRATE (POTENTIAL).
FT	VARIANT	157	224	EGKSLPLDSFSRHLQDOVSLELDSPECTIKYEND PGYMAVCALHPDPEDITVWSHELL -> GROLIASA (IN NON SURFACTIN-PRODUCING STRAINS).
FT	MUTAGEN	105	105	G->A: ALMOST NO ACTIVITY.
FT	MUTAGEN	105	105	G->D: LOSS OF ACTIVITY.
FT	MUTAGEN	107	107	D->E: 3000-FOLD REDUCTION IN ACTIVITY.
FT	MUTAGEN	147	147	BUT NO CHANGE IN SUBSTRATE AFFINITY.
FT	MUTAGEN	147	147	W->A: 24-FOLD REDUCTION IN ACTIVITY.
FT	MUTAGEN	151	151	BUT NO CHANGE IN SUBSTRATE AFFINITY.
FT	MUTAGEN	155	155	W->F: 5-FOLD REDUCTION IN ACTIVITY.
FT	CONFLICT	22	22	BUT NO CHANGE IN SUBSTRATE AFFINITY.
FT	CONFLICT	97	97	E->A: LOSS OF ACTIVITY.
FT	CONFLICT	118	119	K->A: LOSS OF ACTIVITY.
FT	STRAND	2	7	T -> S (IN REF. 2, 3 AND 5).
FT	HELIX	14	21	G -> C (IN REF. 2, 3 AND 5).
FT	HELIX	22	23	IA -> MP (IN REF. 4).
FT	HELIX	26	34	
FT	HELIX	38	58	
FT	TURN	59	60	
FT	HELIX	63	65	
FT	STRAND	69	70	
FT	TURN	72	73	
FT	STRAND	76	77	
FT	TURN	79	80	
FT	TURN	82	83	
FT	STRAND	85	91	
FT	TURN	92	93	
FT	STRAND	94	100	
FT	STRAND	105	110	
FT	HELIX	116	119	
FT	HELIX	125	133	
FT	HELIX	136	157	
FT	TURN	160	161	
FT	STRAND	169	172	
FT	HELIX	174	181	
FT	STRAND	178	181	
FT	TURN	185	186	
FT	STRAND	190	194	
FT	TURN	198	199	
FT	STRAND	200	206	
FT	STRAND	216	217	
FT	HELIX	220	224	
SO	SEQUENCE	224 AA;	26135 MW;	6CBCT6545BBDPE7 CRC64;

Query Match 59.7%; Score 40; DB 1; Length 224;
Best Local Similarity 54.5%; Pred. No. 3.3;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 STDTPGRPLP 12
Db 69 STQYKPCIP 79

RESULT 7
CPDG_CAVPO STANDARD; PRT; 500 AA.

AC 064403; 054866;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D16 (EC 1.14.14.1) (CYP2D16).
GN CYP2D16.
OS Cavia porcellus (Guinea pig).
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Euthera; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1].
RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-38.
RC STRAIN=13; TISSUE=Adrenal cortex;
RX MEDLINE=95251703; PubMed=7733963;
RA Jiang Q., Voigt J.M., Colby H.D.;
RT "Molecular cloning and sequencing of a guinea pig cytochrome P4502D
RT (CYP2D16): high level expression in adrenal microsomes.";
RL Biochem. Biophys. Res. Commun. 209:1149-1156(1995).
RN [2].
RC SEQUENCE FROM N.A.
RC STRAIN=Hartley white; TISSUE=Adrenal gland;
RA Sun Y., Voigt J.M., Pierce J.C., Colby H.D.;
RT "The gene sequence of a xenobiotic metabolism-related cytochrome P450
RT isozyme (CYP2D16) in guinea pig adrenal gland.";
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -I- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -I- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -I- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE INNER ZONE OF
CC THE ADRENAL CORTEX.
CC -I- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U21486; AAA68479.1; -.
DR EMBL; AF020345; AAB94568.1; -.
DR PIR; JC4153; JC4153.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008069; BP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; BP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Microsome; Monooxygenase; Electron transport; Membrane; Heme;
KW Cytochrome; Monooxygenase; Electron transport; Membrane; Heme;
KM Microsome; Endoplasmic reticulum.
FT DOMAIN 81 84 POLY-VAL.
FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 123 123 I -> V (IN REF. 2).
FT CONFLICT 127 127 Y -> N (IN REF. 2).
FT CONFLICT 148 148 G -> R (IN REF. 2).
SQ SEQUENCE 500 AA; 55800 MW; 2429247E49BF6B24 CRC64;
Query Match 59.7%; Score 40; DB 1; Length 500;
Best Local Similarity 72.7%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 STDTPGRPLP 12
Db 98 STDTPDRPLP 108

RESULT 8
UROC_BOVIN STANDARD; PRT; 433 AA.

AC 005589; 028209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

```

DE (U-plasminogen activator).
GN PLAV.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=93216119; PubMed=6385052;
RA Knaetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schlemming W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Inc. Dally J. 5:605-617(1995).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- INDUCTION: By retinoic acid.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L03546; AAA51419.1; -.
DR EMBL; X85801; CA55796.1; -.
DR PIR; JN0560; JN0560.
DR HSSP; P00749; 1LMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRS001144; Urk_plasm_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.

```

```

FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8DADA47027A CRC64;

Query Match 58.2%; Score 39; DB 1; Length 433;
Best Local Similarity 63.6%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASDDTMGPRCL 11
DB 84 AMRDLGRPCL 94

RESULT 9
YHS4 CAEEL STANDARD; PRT; 602 AA.
ID YHS4 CAEEL
AC 018303;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ZK849.4 in chromosome I.
GN ZK849.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderrinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kershaw J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Desfrophin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z82095; CAB05027.2; -.
DR WormPep; ZK849.4; CB25696.
DR InterPro; IPR000615; Desfrophin.
DR Pfam; PF01062; Desfrophin; 1.
DR ProDom; PD002802; Worm_fam_8; 1.
KW Hypothetical protein.
SQ SEQUENCE 602 AA; 67607 MW; 458AB78802BD6B3A CRC64;

Query Match 58.2%; Score 39; DB 1; Length 602;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 DTMGPRCLP 12
DB 416 DTIGKDCLP 424

```


FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 76 76 GAMMA-CARBOXYGLUTAMIC ACID.
 FT DISULFID 61 66 BY SIMILARITY.
 FT DISULFID 91 104 BY SIMILARITY.
 FT DISULFID 109 187 BY SIMILARITY.
 FT DISULFID 130 170 BY SIMILARITY.
 FT DISULFID 158 182 BY SIMILARITY.
 FT DISULFID 215 293 BY SIMILARITY.
 FT DISULFID 236 276 BY SIMILARITY.
 FT DISULFID 264 288 BY SIMILARITY.
 FT DISULFID 333 479 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 388 404 BY SIMILARITY.
 FT DISULFID 533 547 BY SIMILARITY.
 FT DISULFID 561 591 BY SIMILARITY.
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 618 AA; 70268 MW; B89F719A6D601B0 CRC64;

Query Match 58.2%; Score 39; DB 1; Length 618;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 TMGRCLP 12
 Db 231 TLGSPCLP 238

RESULT 11
 PTRR_PIG ID STANDARD; PRT; 585 AA.

AC P50133; 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Parathyroid hormone/parathyroid hormone-related peptide receptor precursor (PTH/PTHr receptor) (PTH/PTHr type I receptor).
 GN PTRR1 OR PTHR.

OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=96823;
 RN [1]

SEQUENCE FROM N.A.
 RA MEDLINE=96305358; PubMed=8688470;
 RA Black E.C., Smith D.P., Zhang X.Y., Frolik C.A., Harvey A.,
 RA Chandrasekhar S., Hailing H.M.,
 RT "Structure and functional expression of a complementary DNA for
 RT porcine parathyroid hormone/parathyroid hormone-related peptide
 RT receptor";
 RL Biochem. Biophys. Acta 1307:339-347(1996).

CC -1- FUNCTION: This is a receptor for parathyroid hormone and for
 CC parathyroid hormone-related peptide. The activity of this receptor
 CC is mediated by G proteins which activate adenylyl cyclase and also
 CC a phosphatidylinositol-calcium second messenger system (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U18315; AAC48619.1; -.
 DR HSSP; 003431; IBLI.
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR001879; hormn_receptor.

DR Pfam; PF00002; 7tm_2; 1.
 DR Pfam; PF02793; HRM; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SM00008; Hormr; 1.
 DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
 DR PROSITE; PS50227; G_PROTEIN_RECP_F2_3; 1.
 DR PROSITE; PS50261; G_PROTEIN_RECP_F2_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 585 PARATHYROID HORMONE/PARATHYROID HORMONE-
 FT RELATED PEPTIDE RECEPTOR.
 FT DOMAIN 27 184 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 185 208 1 (POTENTIAL).
 FT DOMAIN 209 215 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 216 235 2 (POTENTIAL).
 FT DOMAIN 236 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 3 (POTENTIAL).
 FT DOMAIN 302 315 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 316 337 4 (POTENTIAL).
 FT DOMAIN 338 356 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 357 377 5 (POTENTIAL).
 FT DOMAIN 378 404 6 (POTENTIAL).
 FT TRANSMEM 405 423 6 (POTENTIAL).
 FT DOMAIN 424 435 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 436 458 7 (POTENTIAL).
 FT DOMAIN 459 585 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 48 113 BY SIMILARITY.
 FT DISULFID 104 144 BY SIMILARITY.
 FT DISULFID 127 166 BY SIMILARITY.
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 585 AA; 65682 MW; 60BE15CD49B7D210 CRC64;

Query Match 56.7%; Score 38; DB 1; Length 585;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 GRPCLP 12
 Db 101 GRPCLP 106

RESULT 12
 PTRR_MOUSE ID STANDARD; PRT; 591 AA.

AC P41593; 062119;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Parathyroid hormone/parathyroid hormone-related peptide receptor precursor (PTH/PTHr receptor) (PTH/PTHr type I receptor).
 GN PTRR1 OR PTHR.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

SEQUENCE FROM N.A.
 RA STRAIN=C3H/HEHA;
 RA MEDLINE=95034305; PubMed=7524627;
 RA Karpentier M., van Dijk T.B., Hoeljmakers T., Cremers F.,
 RA Abu-Samra A.-B., Boonstra J., de Laat S.W., Defize L.H.K.;
 RT "Expression pattern of parathyroid hormone/parathyroid hormone
 RT related peptide receptor mRNA in mouse postimplantation embryos
 RT indicates involvement in multiple developmental processes."
 RL Mech. Dev. 47:29-42(1994).

CC [2]
 DR SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RC MEDLINE=94255468; PubMed=8197183;
 RX

DR PROSITE; PSS0261; G PROTEIN RECP F2 4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SGNL 1 26 POTENTIAL.
 FT CHAIN 27 591 PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR.
 FT DOMAIN 27 188 1 (POTENTIAL).
 FT TRANSSEM 189 212 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 213 219 2 (POTENTIAL).
 FT TRANSSEM 220 239 2 (POTENTIAL).
 FT TRANSSEM 240 282 3 (POTENTIAL).
 FT TRANSSEM 283 306 3 (POTENTIAL).
 FT TRANSSEM 307 320 4 (POTENTIAL).
 FT TRANSSEM 321 342 4 (POTENTIAL).
 FT TRANSSEM 343 361 5 (POTENTIAL).
 FT TRANSSEM 362 382 5 (POTENTIAL).
 FT TRANSSEM 383 409 6 (POTENTIAL).
 FT TRANSSEM 410 428 6 (POTENTIAL).
 FT TRANSSEM 429 440 7 (POTENTIAL).
 FT TRANSSEM 441 463 7 (POTENTIAL).
 FT TRANSSEM 464 591 7 (POTENTIAL).
 FT TRANSSEM 48 117 7 (POTENTIAL).
 FT TRANSSEM 108 148 7 (POTENTIAL).
 FT TRANSSEM 131 170 7 (POTENTIAL).
 FT TRANSSEM 151 151 7 (POTENTIAL).
 FT TRANSSEM 161 161 7 (POTENTIAL).
 FT TRANSSEM 166 166 7 (POTENTIAL).
 FT TRANSSEM 176 176 7 (POTENTIAL).
 FT TRANSSEM 591 6260 7 (POTENTIAL).
 FT TRANSSEM 21944P3051B9B9C1 CMC64; 7 (POTENTIAL).

Query Match 56.7%; Score 38; DB 1; Length 591;
 Best local similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 GRPCLP 12
 Db 105 GRPCLP 110

RESULT 14

PTNR_HUMAN STANDARD; PRT; 593 AA.
 ID PTNR_HUMAN STANDARD; PRT; 593 AA.
 AC Q03431;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Parathyroid hormone/parathyroid hormone-related peptide receptor precursor (PTH/PTHrP receptor) (PTH/PTHrP type I receptor).
 OS PTHrP OR PTHrP.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=93238641; PubMed=8386612;
 RA Schipani E., Karga H., Karaplis A.C., Potts J.T. Jr., Kronenberg H.M., Abou-Samra A.-B., Segre G.V., Jueppner H.;
 RT "Identical complementary deoxyribonucleic acids encode a human renal and bone parathyroid hormone (PTH)/PTH-related peptide receptor.";
 RL Endocrinology 132:2157-2165 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=93387403; PubMed=8397094;
 RA Schneider H., Feyen U.-H., Rao Movva N.;
 RT "Cloning and functional expression of a human parathyroid hormone receptor.";
 RL Eur. J. Pharmacol. 246:149-155 (1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95263723; PubMed=7745008;
 RA Schipani E., Weinstein L.S., Bergwitz C., Iida-Klein A., Kong X.F.,

RA Stuhmann M., Kruse K., Whyte M.P., Murray T., Schmidtke J., Dop C., Brichman A.S., Crawford J.D., Potts J.T. Jr., Kronenberg H.M., Abou-Samra A.-B., Segre G.V., Jueppner H.;
 RT "Pseudohypoparathyroidism type 1b is not caused by mutations in the coding exons of the human parathyroid hormone (PTH)/PTH-related peptide receptor gene.";
 RL J. Clin. Endocrinol. Metab. 80:1611-1621 (1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Levine M.A.;
 RT "Characterization of cDNA and genomic DNA encoding the human PTH/PTHrP receptor.";
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP EXTRACELLULAR DOMAIN DISULFIDE BONDS.
 RX MEDLINE=20374568; PubMed=10913300;
 RA Graesschopf U., Lilje H., Honold K., Wozny M., Reusch D., Baswein A., Schaefer W., Rucknagel K.P., Rudolph R.;
 RT "The N-terminal fragment of human parathyroid hormone receptor 1 constitutes a hormone binding domain and reveals a distinct disulfide pattern.";
 RL Biochemistry 39:8878-8887 (2000).
 RN [6]
 RP STRUCTURE BY NMR OF 168-198.
 RX MEDLINE=98409426; PubMed=9737850;
 RA Pelligrini M., Bisello A., Rosenblatt M., Chorev M., Mierke D.F.;
 RT "Binding domain of human parathyroid hormone receptor: from conformation to function.";
 RL Biochemistry 37:12737-12743 (1998).
 RN [7]
 RP VARIANT JMC ARG-223.
 RX MEDLINE=95215874; PubMed=7701349;
 RA Schipani E., Kruse K., Jueppner H.;
 RT "A constitutively active mutant PTH-PTHrP receptor in Jansen-type metaphyseal chondrodysplasia.";
 RL Science 268:98-100 (1995).
 RN [8]
 RP VARIANTS JMC ARG-223 AND PRO-410.
 RX MEDLINE=9636745; PubMed=8703170;
 RA Schipani E., Langman C.B., Parfitt A.M., Jensen G.S., Kikuchi S., Koeh S.W., Cole W.G., Jueppner H.;
 RT "Constitutively activated receptors for parathyroid hormone and parathyroid hormone-related peptide in Jansen's metaphyseal chondrodysplasia.";
 RL New Engl. J. Med. 335:708-714 (1996).
 RN [9]
 RP CHARACTERIZATION OF VARIANTS JMC ARG-223 AND PRO-410.
 RX MEDLINE=97322091; PubMed=9178745;
 RA Schipani E., Jensen G.S., Pincus J., Nissen R.A., Gardella T.J., Jueppner H.;
 RT "Constitutive activation of the cyclic adenosine 3',5'-monophosphate signaling pathway by parathyroid hormone (PTH)/PTH-related peptide receptors mutated at the two loci for Jansen's metaphyseal chondrodysplasia.";
 RL Mol. Endocrinol. 11:851-858 (1997).
 RN [10]
 RP VARIANT BOCAL LEO-132.
 RX MEDLINE=98417978; PubMed=9745456;
 RA Zhang P., Jobert A.-S., Couvineau A., Silve C.;
 RT "A homozygous inactivating mutation in the parathyroid hormone/parathyroid hormone-related peptide receptor causing Blomstrand chondrodysplasia.";
 RL J. Clin. Endocrinol. Metab. 83:3365-3368 (1998).
 RN [11]
 RP VARIANT JMC ARG-458.
 RX MEDLINE=99415605; PubMed=10487664;
 RA Schipani E., Langman C.B., Hunzelman J., Le Merrer M., Loke K.Y., Dillon M.J., Silve C., Jueppner H.;
 RT "A novel parathyroid hormone (PTH)/PTH-related peptide receptor mutation in Jansen's metaphyseal chondrodysplasia.";
 RL J. Clin. Endocrinol. Metab. 84:3052-3057 (1999).
 RN [12]

RP VARIANT ENCHONDROMATOSIS CYS-150.
 RX MEDLINE=11918583; PubMed=11850620;
 RA Hopyan S., Gokgoz N., Poon R., Gensure R.C., Yu C., Cole W.G.,
 RA Bell R.S., Jueppner H., Andrusis I.L., Wunder J.S., Alman B.A.;
 RT "A mutant PTHrP type I receptor in enchondromatosis.";
 RL Nat. Genet. 30:306-310(2002).
 CC -1- FUNCTION: This is a receptor for parathyroid hormone and for
 CC parathyroid hormone-related peptide. The activity of this
 CC receptor is mediated by G proteins which activate adenylyl
 CC cyclase and also a phosphatidylinositol-calcium second messenger
 CC system.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in most tissues. Most abundant in
 CC kidney, bone and liver.
 CC -1- DISEASE: Defects in PTHRI are the cause of Jansen's metaphyseal
 CC chondrodysplasia (JMC) [MIM:156400]. JMC is a rare autosomal
 CC dominant disorder characterized by a short-limbed dwarfism
 CC associated with hypercalcemia and normal or low serum
 CC concentrations of the two parathyroid hormones.
 CC -1- DISEASE: Defects in PTHRI are the cause of chondrodysplasia
 CC Blomstrand type (BOCD) [MIM:215045]. BOCD is a severe skeletal
 CC dysplasia.
 CC -1- DISEASE: Defects in PTHRI can be a cause of enchondromatosis
 CC [MIM:166000]. Enchondromas are common benign cartilage tumors of
 CC bone. They can occur as solitary lesions or as multiple lesions in
 CC caused by enchondromas (Ollier and Maffucci diseases). Clinical problems
 CC caused by enchondromas include skeletal deformity and the
 CC potential for malignant change to osteosarcoma.
 CC -1- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04308; AAA56525.1; -;
 DR EMBL; X68596; CAA48589.1; -;
 DR EMBL; U22409; AAB60657.1; -;
 DR EMBL; U22401; AAB60657.1; JOINED.
 DR EMBL; U22402; AAB60657.1; JOINED.
 DR EMBL; U22403; AAB60657.1; JOINED.
 DR EMBL; U22404; AAB60657.1; JOINED.
 DR EMBL; U22405; AAB60657.1; JOINED.
 DR EMBL; U22406; AAB60657.1; JOINED.
 DR EMBL; U22407; AAB60657.1; JOINED.
 DR EMBL; U22408; AAB60657.1; JOINED.
 DR EMBL; U17418; AAA56774.1; -;
 DR PIR; I38139; A49191.
 DR PDB; 1BL1; 30-MAR-89.
 DR PDB; 1ET2; 06-SEP-00.
 DR PDB; 1ET3; 06-SEP-00.
 DR Genew; HGNC:9608; PTHRI.
 DR MIM; 168468; -;
 DR MIM; 156400; -;
 DR MIM; 215045; -;
 DR MIM; 166000; -;
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0004991; F:parathyroid hormone receptor activity; TAS.
 DR GO; GO:0007187; P:G-protein signaling, coupled to cyclic nucl. . . ; TAS.
 DR GO; GO:0001501; P:skeletal development; TAS.
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR001879; hormn_receptor.
 DR Pfam; PFO0002; 7tm_2; 1.
 DR Pfam; PF02793; HRM; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SMD0008; Hormr; 1.
 DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.

DR PROSITE; PS50227; G_PROTEIN_RECP_F2_3; 1.
 DR PROSITE; PS50261; G_PROTEIN_RECP_F2_4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Disease mutation; 3D-structure; Dwarfism.
 FT SIGNAL 1 26
 FT CHAIN 27 593
 FT DOMAIN 27 188
 FT TRANSSEM 189 212
 FT DOMAIN 213 219
 FT TRANSSEM 220 239
 FT DOMAIN 240 282
 FT TRANSSEM 283 306
 FT DOMAIN 307 320
 FT TRANSSEM 321 342
 FT DOMAIN 343 361
 FT TRANSSEM 362 382
 FT DOMAIN 383 409
 FT TRANSSEM 410 428
 FT DOMAIN 429 440
 FT TRANSSEM 441 463
 Query Match 56.7%; Score 38; DB 1; Length 593;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 7 GRPCLP 12
 Db 105 GRPCLP 110
 RESULT 15
 PSB3 TRYB STANDARD; PRT; 205 AA.
 AC QGNDAI;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Proteasome subunit beta type 3 (EC 3.4.25.1) (20S proteasome subunit
 DE beta-3).
 GN PSB3.
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=427;
 RA Bozdech Z., Huang L., Morton A., Wang C.C.;
 RT "bPSB3-beta 3 subunit of 20S proteasome from T. brucei.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: The proteasome is a multicatalytic proteinase complex
 CC which is characterized by its ability to cleave peptides with Arg,
 CC Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or
 CC slightly basic pH. The proteasome has an ATP-dependent proteolytic
 CC activity (By similarity).
 CC -1- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
 CC specificity.
 CC -1- PATHWAY: Involved in an ATP/ubiquitin-dependent non-lysosomal
 CC proteolytic pathway.
 CC -1- SUBUNIT: The proteasome is composed of at least 15 non identical
 CC subunits which form a highly ordered ring-shaped structure (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family T1B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF169653; AAF89685.1; -.
 DR HSSP; P25451; IRYP.
 DR InterPro; IPR000243; Pept_T1A_subB.
 DR InterPro; IPR001353; Peptidase_T1.
 DR Pfam; PF00227; Proteasome; 1.
 DR PROSITE; PS00854; PROTEASOME_B; FALSE NEG.
 KW Proteasome; Hydrolase; Protease; Threonine protease.
 SQ SEQUENCE 205 AA; 22458 MW; 4CB2093E81E2FFD7 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 205;
 Best Local Similarity 54.5%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 STDTPGRPCLP 12
 :||:|
 Db 123 ATDLIGAPCEP 133

Search completed: March 18, 2004, 11:17:01
 Job time : 1.73523 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 3.28228 seconds
(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-4
Perfect score: 67
Sequence: 1 ASTDTWGRPLP 12

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rotent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteria:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	154	4 Q96SE8	Q96SE8 homo sapien
2	51	76.1	231	11 Q8C6L2	Q8C6L2 mus musculi
3	44	65.7	433	6 Q8M1L0	Q8M1L0 oryctolagus
4	44	65.7	433	6 Q8M1L0	Q8M1L0 oryctolagus
5	42	62.7	402	17 Q8TYC9	Q8TYC9 methanopyru
6	40	59.7	165	2 P94401	P94401 bacillus su
7	40	59.7	165	2 Q93UT8	Q93UT8 bacillus su
8	40	59.7	253	2 Q7MX66	Q7MX66 alcaligenes
9	40	59.7	581	16 Q87SD8	Q87SD8 vibrio para
10	40	59.7	608	16 Q8BZW7	Q8BZW7 leprospira
11	39	58.2	787	10 Q96U20	Q96U20 neurospora
12	39	58.2	27	10 Q9S8W4	Q9S8W4 avena sativ
13	39	58.2	157	6 Q9TV48	Q9TV48 bos taurus
14	39	58.2	288	3 Q14136	Q14136 schizosach
15	39	58.2	419	5 Q9VBA5	Q9VBA5 drosophila
16	39	58.2	419	5 Q24548	Q24548 drosophila

17	39	58.2	434	10 Q7XQ04	Q7XQ04 oryza sativ
18	39	58.2	476	3 Q8NJK6	Q8NJK6 aspergillus
19	39	58.2	477	16 Q93J36	Q93J36 streptomyce
20	39	58.2	533	16 Q82NN7	Q82NN7 streptomyce
21	39	58.2	543	16 Q8YY82	Q8YY82 anabaena sp
22	39	58.2	621	10 Q7X619	Q7X619 oryza sativ
23	39	58.2	800	10 Q8S868	Q8S868 oryza sativ
24	39	58.2	800	10 Q7XEL6	Q7XEL6 oryza sativ
25	39	58.2	813	5 Q9BLJ2	Q9BLJ2 ciona intes
26	39	58.2	1063	5 Q8SR27	Q8SR27 encephalito
27	38	56.7	125	16 Q7UMM5	Q7UMM5 rhodospirell
28	38	56.7	125	10 Q9SUZ2	Q9SUZ2 arabidopsis
29	38	56.7	167	17 Q9YAM3	Q9YAM3 aeropyrum p
30	38	56.7	232	16 Q8CUZ6	Q8CUZ6 oceanobacil
31	38	56.7	259	10 Q9ARS4	Q9ARS4 oryza sativ
32	38	56.7	333	4 Q8NSV1	Q8NSV1 homo sapien
33	38	56.7	437	9 Q9AZM2	Q9AZM2 bacterioph
34	38	56.7	437	16 Q9C194	Q9C194 lactococcus
35	38	56.7	540	16 Q9K414	Q9K414 streptomyce
36	38	56.7	589	6 Q9GMD1	Q9GMD1 oryctolagus
37	38	56.7	589	6 Q7YR13	Q7YR13 cervus elap
38	38	56.7	591	11 Q91WV4	Q91WV4 mus musculi
39	38	56.7	591	11 Q80WU8	Q80WU8 mus musculi
40	38	56.7	595	6 Q9TU31	Q9TU31 canis famli
41	38	56.7	618	11 Q99KV8	Q99KV8 mus musculi
42	38	56.7	668	11 Q8R5A3	Q8R5A3 mus musculi
43	38	56.7	688	3 Q87IU3	Q87IU3 neurospora
44	38	56.7	877	10 Q7X103	Q7X103 oryza sativ
45	38	56.7	3956	2 Q7WTF2	Q7WTF2 streptomyce

ALIGNMENTS

RESULT 1
Q96SE8 PRELIMINARY; PRT; 154 AA.
ID Q96SE8
AC Q96SE8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator." in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";
CC Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029537; AK38734.1; -;
DR GO; GO:0016301; F-kinase activity; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.

```

DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF505572 CRC64;

Query Match
Best Local Similarity 100.0%; Score 67; DB 4; Length 154;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
DB 82 ASTDTMGRPCLP 93

RESULT 2
ID Q86GL2 PRELIMINARY; PRT; 231 AA.
AC Q86GL2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
DR PIR; P70534; P70534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle_1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
FT NON TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E980A682737F2 CRC64;

Query Match
Best Local Similarity 76.1%; Score 51; DB 11; Length 231;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCL 11
DB 83 ANTDITGRPCL 93

RESULT 3
ID Q8MIL0 PRELIMINARY; PRT; 433 AA.
AC Q8MIL0;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
OS PLANU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2215945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY122285; AA083187.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR008293; Pept_S1A_tPA.
DR Pfam; PF00051; Kringle_1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR PIRSF; PIRSF001144; Urk plas act; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F3641554980 CRC64;

Query Match
Best Local Similarity 72.7%; Score 44; DB 6; Length 433;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCL 11
DB 84 ANTDIMDRPCL 94

RESULT 4
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
OS UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator. mRNA,
RT complete cds."

```

```

SUBMITTED (JUL-2002) TO THE EMBL/GENBANK/CDSD DATABASES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR GO:0004263; F:chymotrypsin activity; IEA.
DR GO:0016301; F:kinase activity; IEA.
DR GO:0008223; F:peptidase activity; IEA.
DR GO:0004295; F:trypsin activity; IEA.
DR GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR01254; Peptidase_S1.
DR InterPro: IPR01314; Peptidase_S1A.
DR Pfam: IPR008293; Pept_S1A_NPA.
DR Pfam: PF00051; Kringle_1.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR Pfam: PF000395; Kringle_1.
DR SMART: SM00130; KR_1.
DR SMART: SM00020; TRY_SPC_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_2; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR PIRSF: PIRSF001144; Utk_plasm_act; 1.
KM Glycoprotein; Hydrolyase; Kinase; Kringle; Protease; Serine protease.
SO SEQUENCE 433 AA; 48444 MW; 6BD35A371010A6EE CRC64;

Query Match 65.7%; Score 44; DB 6; Length 433;
Best Local Similarity 72.7%; Pred. No. 8.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCL 11
   ||| |||
Db 84 ANTDIMDRPCL 94

RESULT 5
Q8TYC9 PRELIMINARY; PRT; 402 AA.
AC Q8TYC9;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Predicted GPase, probable translation factor.
GN MK0373.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus
CX NCBI_TaxId=2320;
   [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AV19 / DSM 6324 / JCM 9639;
MEDLINE=1927647; PubMed=11930014;
SLategav A.I., Mezheva K.V., Makarova K.S., Polishin N.N.,
Shenberdina O.V., Shaknova V.V., Belova G.I., Aravind L.,
Natalya A.G., Kozgin I.B., Yakushev R.L., Wolf Y.I., Stetter K.O.,
"the complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens".
Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
RL EMBL; AE010334; AA001588.1; -.
DR InterPro: IPR006073; GTP1_OBG.
DR InterPro: IPR004095; TGS_dom.
DR Pfam: PF02824; TGS; 1.
DR PRINTS: PR00326; GTP1OBG.
KM Complete proteome.
SO SEQUENCE 402 AA; 44508 MW; D7D531395D736894 CRC64;

```

```

Query Match          62.7%; Score 42; DB 17; Length 402;
Best Local Similarity 72.7%; Pred. No. 18;
-Matches     8; Conservative      0; Mismatches    3; Indels       0; Gaps        0;

OY           2 STDWGRPCLP 12
Db            116 STDEGRPCDP 126

RESULT 6
P94401 PRELIMINARY; PRT; 165 AA.
ID P94401;
AC P94401;
DT 01-MAY-1997 (TREMBLrel. 03. Created)
DL 01-MAY-1997 (TREMBLrel. 03. Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25. Last annotation update)
DN Surfactin production.
GN SFP.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 tRPC2.
RX MEDLINE=97124189; PubMed=8969502;
RA Yamane K., Kumano M., Kurita K.;
RT "The 25 degree-36 degree region of the Bacillus subtilis chromosome: of
   determination of the sequence of a 146kb segment and identification of
   113 genes."
RL Microbiology 142:3047-3056(1996).
DR EMBL: D50453; BAA08991.1; -.
DR HSSP: P39135; 10R0.
DR InterPro: IPR008278; 4-PPT_transf.
DR Pfam: PF01648; ACPS; 1.
SQ SEQUENCE 165 AA; 19580 MW; 0CSE5001AFD46SDF CRC64;

Query Match          59.7%; Score 40; DB 2; Length 165;
Best Local Similarity 54.5%; Pred. No. 18;
Matches     6; Conservative      2; Mismatches    3; Indels       0; Gaps        0;

OY           2 STDWGRPCLP 12
Db            69 STGEYKPCIP 79

RESULT 7
O93UT8 PRELIMINARY; PRT; 165 AA.
ID O93UT8;
AC O93UT8;
DT 01-DEC-2001 (TREMBLrel. 19. Created)
DL 01-DEC-2001 (TREMBLrel. 19. Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25. Last annotation update)
DN Inactive surfactin production protein.
GN SFP-0.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Roongawang N., Morikawa M., Kanaya S.;
RT "sfp-0 gene of a surfactin nonproducing Bacillus subtilis Wllj.".
DL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AB061356; BAB47573.1; -.
DR InterPro: IPR008278; 4-PPT_transf.
DR Pfam: PF01648; ACPS; 1.
SQ SEQUENCE 165 AA; 19568 MW; 6BAE919A917BICA CRC64;

Query Match          59.7%; Score 40; DB 2; Length 165;
Best Local Similarity 54.5%; Pred. No. 18;
Matches     6; Conservative      2; Mismatches    3; Indels       0; Gaps        0;

OY           2 STDWGRPCLP 12
```

Db 69 STGEYKPCIP 79

RESULT 8

Q7WX66 PRELIMINARY; PRT; 253 AA.
AC Q7WX66;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Transcriptional regulator.
GN FNR OR PHG276.
OS Alkaligenes eutrophus (Ralstonia eutropha).
OC Plasmid megaplasmid PHG1.
CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
CC Burkholderiaceae; Ralstonia.
CX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H16;
RA Schwartz E., Henne A., Cramm R., Eitinger T., Friedrich B.,
RT "Complete Nucleotide Sequence of pHG1: A Ralstonia eutropha H16
RT Megaplasmid Encoding Key Enzymes of H2-based lithoautotrophy and
RT Anaerobiosis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY305378; AAP86025.1; -.

QY 253 AA; 28316 MW; P664DD75198CE1E9 CRC64;
SQ SEQUENCE

Query Match 59.7%; Score 40; DB 2; Length 253;
Best Local Similarity 72.7%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 27 STCTMGRCPLP 12
Db 27 STCTMGRCPLP 37

RESULT 9

Q87SD8 PRELIMINARY; PRT; 581 AA.
AC Q87SD8;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Sensory box/GDEF family protein.
GN VP0486.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
CX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; Pubmed=12620739;
RA Matino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005074; BACS8749.1; -.

DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; F:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001610; GDEF.
DR InterPro; IPR000700; PAS-aassoc_C.
DR Pfam; PF00990; GDEF. 1.
DR Pfam; PF00785; PAC. 1.
DR PROSITE; PS50887; GDEF. 1.
DR PROSITE; PS50113; PAC. 2.

KW Complete proteome.
SQ SEQUENCE 581 AA; 66107 MW; 1830D37B51E24994 CRC64;

Query Match 59.7%; Score 40; DB 16; Length 581;
Best Local Similarity 70.0%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 216 TTDTQGRSCL 225
Db 216 TTDTQGRSCL 225

RESULT 10

Q8EZM7 PRELIMINARY; PRT; 608 AA.
AC Q8EZM7;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN LA3734.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
CX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar 1a1;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011529; AAN50932.1; -.

QY 608 AA; 63971 MW; DA173AAB9FC21B1F CRC64;
SQ SEQUENCE

Query Match 59.7%; Score 40; DB 16; Length 608;
Best Local Similarity 58.3%; Pred. No. 64;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASTDIMGRCPLP 12
Db 589 AATDILGRPLP 600

RESULT 11

Q96U20 PRELIMINARY; PRT; 787 AA.
AC Q96U20;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN B208.430.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
CX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hohnsels J., Brandt P., Fartmann B., Holland R.,
RA Wakamura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL155930; CAD01134.1; -.
DR GO; GO:0003743; P:translation initiation factor activity; IEA.
DR GO; GO:0006413; P:translational initiation; IEA.
DR InterPro; IPR001950; TIF_SUI1.
DR PROSITE; PS01118; SUI1_1; 1.
KW Hypothetical protein
SQ SEQUENCE 787 AA; 84646 MW; BCAL1937ABEB5B3 CRC64;

Query Match 59.7%; Score 40; DB 3; Length 787;

Best Local Similarity 77.8%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASTDTWGRP 9
Db 219 SSTDSMGRP 227

RESULT 12

O9S8M4 PRELIMINARY; PRT; 27 AA.
AC O9S8M4; 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Alpha-amylase inhibitor (Fragment).
OS Avena sativa (Oat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Avenae; Avena.
OX NCBI_TaxID=4498;
RN [1]
RP SEQUENCE
RA MEDLINE=92405739; Pubmed=1526282;
RA Roeder A., Collila F., Ortiz M.L., Mendez E.;
RT "Identification of the three major coeliac immunoreactive proteins and
RT one alpha-amylase inhibitor from oat endosperm."
RL PIR, S29211, S29211.
DR HSP, P07597, 1BE2.
SQ SEQUENCE 27 AA; 2756 MW; E48BD7DA87385341 CRC64;

Query Match 58.2%; Score 39; DB 10; Length 27;
Best Local Similarity 66.7%; Pred. No. 4.5;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 DTWGRPLP 12
Db 7 DTLVKRCLP 15

RESULT 13

O9TVAB PRELIMINARY; PRT; 157 AA.
AC O9TVAB; 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071386; Pubmed=11204721;
RA Balcerzak D., Querengesser U., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle."
RL J. Anim. Sci. 79:94-107(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL, AF144761; AAD30301.1; -.
DR HSP, P00749, 1URK.
DR GO, GO:0016301, F:kinase activity, IEA.
DR InterPro, IPR006209; EGF_1ike.
DR InterPro, IPR000001; Kringle.
DR Pfam, PF00051; kringle.1.
DR PRINTS, PRO0018; KRINGLE.
DR ProDom, PD000395; Kringle.1.
DR SMART, SM00130; KR; 1.
DR PROSITE, PS00022; EGF_1; 1.

DR PROSITE, PS00021; KRINGLE 1; 1.
DR PROSITE, PS50070; KRINGLE 2; 1.
KM Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7.CRC64;

Query Match 58.2%; Score 39; DB 6; Length 157;
Best Local Similarity 63.6%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ASTDTWGRP 11
Db 50 ANRDLGRPCL 60

RESULT 14

O14136 PRELIMINARY; PRT; 288 AA.
AC O14136;
DT 01-JAN-1999 (TReMBLrel. 09, Created)
DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Putative peroxisomal membrane protein, with SRC homology domain.
GN SPAC3C7.10.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Brown D., Churcher C.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Wood V., Bartell B.G., Rajandream M.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL, Z99568; CAB16740.1; -.
DR PIR, T38696; T38696.
DR GenBank, SPombe; SPAC3C7.10; -.
DR InterPro, IPR007223; Peroxin-13_N.
DR InterPro, IPR001452; SH3.
DR Pfam, PF04088; Peroxin-13_N; 1.
DR Pfam, PF00018; SH3; 1.
DR PRINTS, PRO0452; SH3DOMAIN.
DR ProDom, PD000066; SH3; 1.
DR SMART, SM00326; SH3; 1.
DR PROSITE, PS50002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 288 AA; 32054 MW; 992FP93538CDE1B43 CRC64;

Query Match 58.2%; Score 39; DB 3; Length 288;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ASTDTWGRP 10
Db 254 SKDTQGNPC 263

RESULT 15

O9VBAS PRELIMINARY; PRT; 419 AA.
AC O9VBAS;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE T48 protein (GMI8993p).
GN T48 OR CG5507.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 4.13567 Seconds

(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-5

Perfect score: 62

Sequence: 1 RRPWCYQVQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	10	5	AAG79463
2	62	100.0	411	2	AAV39343
3	62	100.0	411	2	AAV42284
4	62	100.0	411	5	AAE13269
5	62	100.0	411	7	ADC72159
6	57	91.9	65	1	AAV93232
7	57	91.9	86	7	ABR42617
8	57	91.9	86	7	ABR42602
9	57	91.9	86	7	ABR42599
10	57	91.9	86	7	ABR42601
11	57	91.9	87	7	ABR42604
12	57	91.9	87	7	ABR42605
13	57	91.9	88	5	AAE16542
14	57	91.9	89	2	AAW22742
15	57	91.9	96	5	AAE16550
16	57	91.9	135	5	AAE16545
17	57	91.9	143	5	AAE16549
18	57	91.9	194	5	AAW22745
19	57	91.9	201	2	AAW22746
20	57	91.9	322	7	ABR42615
21	57	91.9	322	7	ABR42616
22	57	91.9	337	5	AAV75492
23	57	91.9	337	5	ABP41795
24	57	91.9	365	2	AAE68854
25	57	91.9	378	2	AAW13635

26	57	91.9	386	2	AAE66266	AAE66266	Bifunctio
27	57	91.9	389	2	AAW13636	AAW13636	Human pro
28	57	91.9	390	2	AAE66245	AAE66245	Bifunctio
29	57	91.9	390	2	AAE66247	AAE66247	Bifunctio
30	57	91.9	392	2	AAE66260	AAE66260	Bifunctio
31	57	91.9	392	2	AAE66264	AAE66264	Bifunctio
32	57	91.9	392	2	AAE66255	AAE66255	Bifunctio
33	57	91.9	392	2	AAE66259	AAE66259	Bifunctio
34	57	91.9	392	2	AAE66261	AAE66261	Bifunctio
35	57	91.9	392	2	AAE66258	AAE66258	Bifunctio
36	57	91.9	392	2	AAE66263	AAE66263	Bifunctio
37	57	91.9	392	2	AAE66254	AAE66254	Bifunctio
38	57	91.9	392	2	AAE66256	AAE66256	Bifunctio
39	57	91.9	392	2	AAE66257	AAE66257	Bifunctio
40	57	91.9	393	2	AAE66251	AAE66251	Bifunctio
41	57	91.9	393	2	AAE66253	AAE66253	Bifunctio
42	57	91.9	393	2	AAE66249	AAE66249	Bifunctio
43	57	91.9	393	2	AAE66252	AAE66252	Bifunctio
44	57	91.9	393	2	AAE66244	AAE66244	Bifunctio
45	57	91.9	393	2	AAE66250	AAE66250	Bifunctio

ALIGNMENTS

RESULT 1
AAG79463
ID AAG79463 standard; peptide; 10 AA.
XX
AC AAG79463;
XX
DT 15-NOV-2002 (first entry)
XX
DE sc-uPA long A chain peptide fragment 3, amino acids 129-138.
XX
KW Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
KW high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KW long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KW low molecular weight urokinase-type plasminogen activator; LMW-uPA; CD87;
KW binding domain.
XX
OS Homo sapiens.
XX
PN EP1232755-A2.
XX
PD 21-AUG-2002.
XX
PF 15-FEB-2002; 2002EP-00003555.
XX
PR 20-FEB-2001; 2001JP-00042655.
PR 19-JUN-2001; 2001JP-00184284.
XX
PA (JCRP-) JCR PHARM CO LTD.
XX
PI Wada M, Wada N;
XX
XX WPI; 2002-610512/66.
DR
XX
PT Anti-HIV agents; comprises ligand molecule that binds to CD87, e.g. high
PT molecular weight urokinase-type plasminogen activator, amino-terminal
PT fragment or an anti-CD87 antibody.
XX
XX
PS Example; Page 23; 38pp; English.
XX
XX The sequences given in AAG79461-63 represent peptide fragments derived
XX from the long A chain of single chain prepro-urokinase (sc-uPA). Pro-
XX urokinase (amino acids 21-431) with a cleavage between amino acids 178
XX and 179 gives high molecular weight urokinase-type plasminogen activator
XX (HMW-uPA). HMW-uPA is a protein consisting of two peptide chains linked
XX by a di-sulphide bond. The chains, long A and B, are formed by enzymatic
XX cleavage between amino acids 178 and 179 of pro-urokinase. HMW-uPA
XX includes an EGF-like domain, a kringle domain and a urokinase receptor
XX (CD87) binding domain. HMW-uPA is then cleaved between amino acids 155

CC and 156 to give low molecular weight urokinase-type plasminogen activator
 CC (LMW-UPA) (amino acids 156-178 and 179-431), that has no plasminogen
 CC activator activity. sc-UPA, or these fragments of it may be used in the
 CC anti-HIV agents of the invention which comprise a ligand molecule that
 CC binds to CD87. The agents are useful for treating HIV-infected humans for
 CC suppression of reproduction of HIV. The anti-HIV agents act by a
 CC mechanism of action different from those of conventional drugs, widening
 CC the choice of therapeutics agents and avoiding problems of resistant HIV
 XX

Sequence 10 AA;

Query Match 100.0%; Score 62; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0072;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQVQ 10
 Db 1 RRPWCYQVQ 10

RESULT 2

AA39343

AA39343 standard; protein; 411 AA.

01-DEC-1999 (first entry)

Human pro-urokinase.

Serine protease; plasminogen; plasmin; activation; matrix; cancer;
 tumour; metastasis; X-ray crystallography; inhibitor.

Homo sapiens.

Location/Qualifiers

Key 1. .158

Protein /label= Mature_urokinase_A_chain

Protein /label= Low_molecular_weight_urokinase

Disulfide-bond 11. .19

Disulfide-bond 13. .31

Disulfide-bond 33. .42

Disulfide-bond 50. .131

Disulfide-bond 71. .113

Disulfide-bond 102. .126

Cleavage-site 135. .136

/note= "Additional cleavage at this site generates low
 molecular weight (LMW) urokinase"

Disulfide-bond 148. .279

/note= "Links mature urokinase A- and B-chains"

Cleavage-site 158. .159

/note= "Cleavage at this site generates mature urokinase
 A- and B-chains"

Protein /label= Mature_urokinase_B_chain

Disulfide-bond 189. .205

Disulfide-bond 197. .268

Disulfide-bond 293. .362

Modified-site 302

/note= "N-glycosylated"

Disulfide-bond 325. .341

Disulfide-bond 352. .380

Cleavage-site 405. .406

WO9945379-A2.

10-SEP-1999.

05-MAR-1999; 99WO-US004967.

06-MAR-1998; 98US-00036184.

PA (ABBO) ABBOTT LAB.

PI Member VL, Greer J, Abad-Zapatero C, Norbeck DW,

XX WPI, 1999-571607/48.

DR Identifying ligands for target biomolecules using X-ray crystallography.

PT Example 1; Fig 5; 57pp; English.

PS This sequence represents human pro-urokinase. The mature urokinase

CC consists of an A- and B-chain, linked by a single disulfide bond, and is

CC generated by proteolytic cleavage of the peptide bond between Lys 158 and

CC Ile 159. Additional cleavage of the peptide bond between Lys 135 and Lys

CC 136 generates a low molecular weight urokinase. The urokinase A-chain

CC contains an EGF-like domain and a kringle domain, while the B-chain

CC contains the catalytic domain. Urokinase is a serine protease and is

CC strongly associated with tumour cells. Urokinase activates plasminogen

CC which, in turn, activates the matrix metalloproteinases. Plasmin and the

CC metalloproteinases degrade the extracellular matrix and promote tumour

CC growth and metastasis. Inhibitors that specifically target urokinase may

CC serve as effective anticancer agents. A novel method for identifying such

CC ligands used X-ray crystallography to determine if a complex is formed

CC between a ligand and a target biomolecule. However, crystals of a native

CC urokinase/inhibitor complex had poor diffraction quality. Human urokinase

CC was therefore engineered so that it would produce crystals with the

CC desired qualities. This engineered urokinase was designated mu-UK

CC (AA39344)

Sequence 411 AA;

Query Match 100.0%; Score 62; DB 2; Length 411;

Best Local Similarity 100.0%; Pred. No. 0.19;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQVQ 10

Db 109 RRPWCYQVQ 118

RESULT 3

AA42284

AA42284 standard; protein; 411 AA.

01-DEC-1999 (first entry)

Human pro-urokinase.

Serine protease; plasminogen; plasmin; activation; matrix; cancer;
 tumour; metastasis; X-ray crystallography; inhibitor.

Homo sapiens.

Location/Qualifiers

Key 1. .158

Protein /label= Mature_urokinase_A_chain

Protein /label= Low_molecular_weight_urokinase

Disulfide-bond 11. .19

Disulfide-bond 13. .31

Disulfide-bond 33. .42

Disulfide-bond 50. .131

Disulfide-bond 71. .113

Disulfide-bond 102. .126

Cleavage-site 135. .136

/note= "Additional cleavage at this site generates low
 molecular weight (LMW) urokinase"

Disulfide-bond 148. .279

/note= "Links mature urokinase A- and B-chains"

Cleavage-site 158. .159

/note= "Cleavage at this site generates mature urokinase

AC ADC72159;
 XX
 DE 18-DEC-2003 (first entry)
 XX
 DE Human urokinase mature protein sequence related to ligand identification.
 XX
 KW ligand; target biomolecule; crystal; X-ray crystal diffraction pattern;
 KW structural change; binding event; crystallographic data;
 KW three dimensional electron density; ligand-receptor complex; human;
 KW urokinase; enzyme.
 XX
 OS Homo sapiens.
 XX
 FN US2003049678-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 07-AUG-2001; 2001US-00923830.
 XX
 PR 05-MAR-1999; 99US-00263904.
 XX
 PA (NIEN/) NIENBER V L.
 PA (GREB/) GREER J.
 PA (ABAD/) ABAD-ZAPATERO C.
 PA (NORB/) NORBECK D W.
 XX
 PI Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;
 DR WPI, 2003-786861/74.
 XX
 PT Identification or designing of ligand to target biomolecule by exposing
 PT target biomolecule crystal to test sample(s), and obtaining X-ray crystal
 PT diffraction pattern to determine, whether ligand/receptor complex is
 PT formed.
 XX
 PS Disclosure; Fig 5; 33pp; English.
 XX
 CC This invention relates to a novel method for the identification of
 CC ligands to a target biomolecule identified or designed by obtaining a
 CC target biomolecule crystal, exposing the target biomolecule crystal to a
 CC test sample and obtaining X-ray crystal diffraction pattern to determine
 CC whether a ligand-receptor complex is formed. The method is useful for
 CC identifying or designing a ligand to a target biomolecule. The invention
 CC provides direct identification of potential ligands and detailed
 CC information on how the ligand binds and changes in the target
 CC biomolecule. It requires non special labelling of the target. It is
 CC uniquely sensitive to structural changes in the target and the ligand.
 CC The binding event is monitored directly, so that the probability for
 CC false positives is reduced to near zero. The crystallographic data
 CC provide a three dimensional electron density snap-shot of the ligand-
 CC receptor complex showing, which compound binds and how it is bound. The
 CC present sequence is the amino acid sequence of human urokinase which was
 CC used during the exemplification of the invention.
 CC
 SQ Sequence 411 AA;
 QY
 Db 1 RRPWCYVQV 10
 109 RRPWCYVQV 118
 100.0%; Score 62; DB 7; Length 411;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 6
 AAP93232
 ID AAP93232 standard; peptide; 65 AA.
 XX
 AC AAP93232;
 XX
 DT 25-MAR-2003 (revised)
 DT 03-APR-1990 (first entry)

XX
 DE Modified scu-PA kringle domain (residues 66-131).
 DE
 KW plasminogen; activator; t-PA; fibrin; kringle domain;
 KW single-chain urinary plasminogen activator; scu-PA.
 XX
 OS Homo sapiens.
 XX
 FN WO8910401-A.
 XX
 PD 02-NOV-1989.
 XX
 PF 22-APR-1988; 88US-00184823.
 XX
 PR 22-APR-1988; 88US-00184823.
 XX
 PA (COLB) COLLABORATIVE RES INC.
 XX
 PI Mao J, Abercromb1 DW;
 XX
 DR WPI, 1989-339965/46.
 DR N-PADB; AAN92239.
 XX
 PT Modified plasminogen activator - having greater fibrin selectivity and
 PT circulating half life.
 XX
 PS Disclosure; Fig 7; 80pp; English.
 XX
 CC The sequence is encoded by a synthetic oligonucleotide and is a modified
 CC scu-PA kringle domain. The modified scu-PA has greater fibrin selectivity
 CC and a longer half-life than normal scu-PA. The scu-PA is used in the
 CC treatment of pulmonary embolism, thrombosis, myocardial infarct and
 CC strokes. (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 65 AA;
 QY
 Db 1 RRPWCYVQV 9
 43 RRPWCYVQV 51
 91.9%; Score 57; DB 1; Length 65;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 7
 ABR42617
 ID ABR42617 standard; protein; 86 AA.
 AC ABR42617;
 XX
 DT 26-AUG-2003 (first entry)
 XX
 DE Human abrogen (hATF-kringle).
 XX
 KW Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
 KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 FN Key Location/Qualifiers
 FN Misc-difference 43
 FT /label= Aasn, Asp
 FT Misc-difference 74
 FT /label= Pro, Leu
 XX
 PN WO2003042354-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 04-SEP-2002; 2002WO-US027885.

04-SEP-2001; 2001US-0316300P.
(AVET) AVENTIS PHARM INC.
Nesbitt M, Fong TC, Brockstedt D;
WPI; 2003-449566/42.
N-PSDB; ACCS8338.
New abrogen polypeptide, useful for treating an angiogenesis related diseases e.g. tumor metastasis.
Disclousure; Page 95; 95pp; English.
The present sequence is the protein sequence of a novel human abrogen, designated hATF-kringle, comprising the human urokinase plasminogen activator kringle domain. Abrogens such as hATF-kringle are potent inhibitors of endothelial proliferation and angiogenesis. Abrogen polypeptides are capable of inhibiting or reducing cell proliferation induced by both basic fibroblast growth factor (bFGF) and vascular endothelial growth factor in a specific endothelial cell proliferation assay; angiostatin only inhibits bFGF induced proliferation in this assay. Vectors that expressed abrogen polypeptides in vivo were shown to reduce tumour metastasis in 2 lung cancer models. The invention provides abrogen polypeptides and polynucleotides, and methods of using these to treat an angiogenesis-related disease or disorder, e.g. tumour metastasis (claimed)

Sequence 86 AA:
SQ

Query Match 91.9%; Score 57; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 RRPWCYQV 9
62 RRPWCYQV 70

RESULT 8
ABR42602
ID ABR42602 standard; protein; 86 AA.
AC ABR42602;
DT 26-AUG-2003 (first entry)
DE Mouse abrogen (hATF-kringle).
KM Mouse; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KM urokinase plasminogen activator; tumour; metastasis; cyostatic;
KM gene therapy.
OS Mus sp.
FH Key Location/Qualifiers
FT Domain 53..59 /note="kringle domain"
WO2003042354-A2.
22-MAY-2003.
04-SEP-2002; 2002MO-US027885.
04-SEP-2001; 2001US-0316300P.
(AVET) AVENTIS PHARM INC.
Nesbitt M, Fong TC, Brockstedt D;
WPI; 2003-449566/42.
N-PSDB; ACCS8337.

New abrogen polypeptide, useful for treating an angiogenesis related diseases e.g. tumor metastasis.
Claim 1; Page 25; 95pp; English.
The present sequence is the protein sequence of a novel mouse abrogen, designated mATF-kringle, comprising the human urokinase plasminogen activator kringle domain. Abrogens such as mATF-kringle are potent inhibitors of endothelial proliferation and angiogenesis. Abrogen polypeptides are capable of inhibiting or reducing cell proliferation induced by both basic fibroblast growth factor (bFGF) and vascular endothelial growth factor in a specific endothelial cell proliferation assay; angiostatin only inhibits bFGF induced proliferation in this assay. Vectors that expressed abrogen polypeptides in vivo were shown to reduce tumour metastasis in 2 lung cancer models. The invention provides abrogen polypeptides and polynucleotides, and methods of using these to treat an angiogenesis-related disease or disorder, e.g. tumour metastasis (claimed)

Sequence 86 AA;
SQ

Query Match 91.9%; Score 57; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 RRPWCYQV 9
62 RRPWCYQV 70

RESULT 9
ABR42599
ID ABR42599 standard; protein; 86 AA.
AC ABR42599;
DT 26-AUG-2003 (first entry)
DE Human abrogen (hATF-kringle).
KM Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KM urokinase plasminogen activator; tumour; metastasis; cyostatic;
KM gene therapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 53..59 /note="kringle domain"
WO2003042354-A2.
22-MAY-2003.
04-SEP-2002; 2002MO-US027885.
04-SEP-2001; 2001US-0316300P.
(AVET) AVENTIS PHARM INC.
Nesbitt M, Fong TC, Brockstedt D;
WPI; 2003-449566/42.
N-PSDB; ACCS8334.
New abrogen polypeptide, useful for treating an angiogenesis related diseases e.g. tumor metastasis.
Claim 1; Page 24; 95pp; English.
The present sequence is the protein sequence of a novel human abrogen, designated hATF-kringle, comprising the human urokinase plasminogen

CC activator kringle domain. Abrogens such as hATF-kringle are potent
 CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
 CC polypeptides are capable of inhibiting or reducing cell proliferation
 CC induced by both basic fibroblast growth factor (bFGF) and vascular
 CC endothelial growth factor in a specific endothelial cell proliferation
 CC assay; angiostatin only inhibits bFGF induced proliferation in this
 CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
 CC reduce tumour metastasis in 2 lung cancer models. The invention provides
 CC abrogen polypeptides and polynucleotides, and methods of using these to
 CC treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
 CC (claimed)

CC Sequence 86 AA;

Query Match 91.9%; Score 57; DB 7; Length 86;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYVQV 9
 |||||
 Db 62 RRPWCYVQV 70

RESULT 10

ABR42601
 ID ABR42601 standard; protein; 86 AA.

AC ABR42601;

DT 26-AUG-2003 (first entry)

DE Human abrogen (hATF-kringle).

Human; abrogen; kringler; hATF-kringle; angiogenesis; inhibitor;
 KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
 KW gene therapy.

OS Homo sapiens.

Key Location/Qualifiers
 FT 53..59
 FT Domain /note="kringle domain"

WO2003042354-A2.

PD 22-MAY-2003.

PF 04-SEP-2002; 2002WO-US027885.

PR 04-SEP-2001; 2001US-0316300P.

PA (AVET) AVENTIS PHARM INC.

PI Nesbit M, Fong TC, Brockstedt D;

DR WPI; 2003-449566/42.

DR N-PSDB; ACC58336.

PT New abrogen polypeptide, useful for treating an angiogenesis related
 PT diseases e.g. tumor metastasis.

PS Claim 1; Page 25; 95pp; English.

CC The present sequence is the protein sequence of a novel human abrogen,
 CC designated hATF-kringle, comprising the human urokinase plasminogen
 CC activator kringler domain. Abrogens such as hATF-kringle are potent
 CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
 CC polypeptides are capable of inhibiting or reducing cell proliferation
 CC induced by both basic fibroblast growth factor (bFGF) and vascular
 CC endothelial growth factor in a specific endothelial cell proliferation
 CC assay; angiostatin only inhibits bFGF induced proliferation in this
 CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
 CC reduce tumour metastasis in 2 lung cancer models. The invention provides

CC abrogen polypeptides and polynucleotides, and methods of using these to
 CC treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
 CC (claimed)

CC Sequence 86 AA;

Query Match 91.9%; Score 57; DB 7; Length 86;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYVQV 9
 |||||
 Db 62 RRPWCYVQV 70

RESULT 11

ABR42604
 ID ABR42604 standard; protein; 87 AA.

AC ABR42604;

DT 26-AUG-2003 (first entry)

DE Human abrogen (hATF-kringle).

Human; abrogen; kringler; hATF-kringle; angiogenesis; inhibitor;
 KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
 KW gene therapy.

OS Homo sapiens.

Key Location/Qualifiers
 FT 54..60
 FT Domain /note="kringle domain"

WO2003042354-A2.

PD 22-MAY-2003.

PF 04-SEP-2002; 2002WO-US027885.

PR 04-SEP-2001; 2001US-0316300P.

PA (AVET) AVENTIS PHARM INC.

PI Nesbit M, Fong TC, Brockstedt D;

DR WPI; 2003-449566/42.

PT New abrogen polypeptide, useful for treating an angiogenesis related
 PT diseases e.g. tumor metastasis.

PS Claim 37; Page 26; 95pp; English.

CC The present sequence is the protein sequence of a secreted form of a
 CC novel human abrogen, designated hATF-kringle, comprising the human
 CC urokinase plasminogen activator kringler domain. The sequence includes an
 CC N-terminal alanine residue that results from cleavage of an interleukin-2
 CC signal peptide. Abrogens such as hATF-kringle are potent inhibitors of
 CC endothelial proliferation and angiogenesis. Abrogen polypeptides are
 CC capable of inhibiting or reducing cell proliferation induced by both
 CC basic fibroblast growth factor (bFGF) and vascular endothelial growth
 CC factor in a specific endothelial cell proliferation assay; angiostatin
 CC only inhibits bFGF induced proliferation in this assay. Vectors that
 CC expressed abrogen polypeptides in vivo were shown to reduce tumour
 CC metastasis in 2 lung cancer models. The invention provides abrogen
 CC polypeptides and polynucleotides, and methods of using these to treat an
 CC angiogenesis-related disease or disorder, e.g. tumour metastasis
 CC (claimed)

CC Sequence 87 AA;

Query Match 91.9%; Score 57; DB 7; Length 87;

Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 63 RRPWCYQV 71

RESULT 12

ABR42605 standard; protein; 87 AA.

AC ABR42605;
DT 26-AUG-2003 (first entry)
XX
DE Human abrogen (hATF-kringle).

KM Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
XX gene therapy.

OS Homo sapiens.

Key Location/Qualifiers
FT Domain 54..60
FT /note= "kringle domain"

PN MO2003042354-A2.

PD 22-MAY-2003.

PF 04-SEP-2002; 2002MO-US027885.

PR 04-SEP-2001; 2001US-0316300P.

PA (AVET) AVENTIS PHARM INC.

PI Nesbit M, Fong TC, Brocketed D;

XX WPI; 2003-449566/42.

DR New abrogen polypeptide, useful for treating an angiogenesis related
PT diseases e.g. tumor metastasis.

PS Claim 37; Page 26; 95pp; English.

XX The present sequence is the protein sequence of a secreted form of a
CC novel human abrogen, designated hATF-kringle, comprising the human
CC urokinase plasminogen activator kringle domain. The sequence includes an
CC N-terminal alanine residue that results from cleavage of an interleukin-2
CC signal peptide. Abrogens such as hATF-kringle are potent inhibitors of
CC endothelial proliferation and angiogenesis. Abrogen polypeptides are
CC capable of inhibiting or reducing cell proliferation induced by both
CC basic fibroblast growth factor (bFGF) and vascular endothelial growth
CC factor in a specific endothelial cell proliferation assay; angiostatin
CC only inhibits bFGF induced proliferation in this assay. Vectors that
CC expressed abrogen polypeptides in vivo were shown to reduce tumour
CC metastasis in 2 lung cancer models. The invention provides abrogen
CC polypeptides and polynucleotides, and methods of using these to treat an
CC angiogenesis-related disease or disorder, e.g. tumour metastasis
CC (claimed)
XX

SO Sequence 87 AA;

Query Match 91.9%; Score 57; DB 7; Length 87;

Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 63 RRPWCYQV 71

RESULT 13

AAE16542 standard; protein; 88 AA.

AC AAE16542;
DT 09-APR-2002 (first entry)
XX

DE Human urokinase-type plasminogen activator (uPA) kringle.

KM Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW male impotence; adult respiratory distress syndrome.

OS Homo sapiens.

PN MO20019752-A2.

PD 27-DEC-2001.

PF 13-JUN-2001; 2001MO-US018976.

PR 20-JUN-2000; 2000US-0212874P.

PA (TYPE-) UNIV PENNSYLVANIA.

PI Cines DB, Higazi AA;

DR WPI; 2002-122240/16.

DN N-PSDB; AAD27075.

PT Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator.

PS Claim 1; Fig 1A; 117pp; English.

XX The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) kringle
XX

SO Sequence 88 AA;

Query Match 91.9%; Score 57; DB 5; Length 88;

Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 62 RRPWCYQV 70

RESULT 14

AAW22742 standard; protein; 89 AA.

ID AAW22742
AC AAW22742;
XX

DT 12-MAR-1998 (first entry)
 XX
 DE Urokinase residues 43-131.
 XX
 KW Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
 KM G-domain; urokinase; prevention; leukaemia; lymphoma.
 XX
 OS Homo sapiens.
 XX
 PN MO9725422-A1.
 PD 17-JUL-1997.
 XX
 PF 06-JAN-1997; 97WO-JP000008.
 XX
 PR 08-JAN-1996; 96JP-00001059.
 XX
 PA (NISP) NISSIN FOOD PROD CO LTD.
 PI Kobayashi H, Terao T, Sugino D, Okushima M;
 XX
 DR WPI; 1997-372862/34.
 XX
 PT Chimeric protein which inhibits development of metastases in cancer -
 PT contains urinary trypsin inhibitor carboxy-terminal domain linked to
 XX urokinase G-domain.
 XX
 PS Claim 3; Page 72; 97pp; Japanese.
 XX
 CC A novel chimeric protein contains the carboxy-terminal domain of human
 CC urinary trypsin inhibitor (HI-8), which inhibits cancer cell metastasis,
 CC linked to a peptide containing the G-domain of urokinase (AAW22742),
 CC which specifically binds the excess urokinase receptor expressed in
 CC cancer cells. The chimeric protein has the amino-terminal AAW22734, the
 CC carboxy-terminal AAW22735 and a linking sequence selected from AAW22736-
 CC 39 or partial sequences derived from these, specifically AAW38130-63. The
 CC chimeric protein may also have additional amino-terminal sequences
 CC selected from AAW22740 or 9 partial sequences derived from this, and/or
 CC additional carboxy-terminal sequences selected from AAW22743 or 10
 CC partial sequences derived from this. The chimeric protein can be used to
 CC prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
 CC stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
 CC prostate, and in leukaemia or lymphoma
 CC
 XX
 SQ Sequence 89 AA;
 QY
 DB 1 RRPWCYQV 9
 67 RRPWCYQV 75
 91.9%; Score 57; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 15
 ID AAE16550
 AAE16550 standard; protein; 96 AA.
 AC AAE16550;
 DT 09-APR-2002 (first entry)
 DE Human uPA kringle and connecting peptide.
 KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; male impotence.
 KM
 XX
 OS Homo sapiens.

XX
 PN MO200197752-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 13-JUN-2001, 2001MO-US018976.
 XX
 PR 20-JUN-2000; 2000US-0212874P.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 PI Cines DB, Higazi AA;
 XX
 DR WPI; 2002-122240/16.
 XX
 DR N-PSDB; AAD27083.
 XX
 PT Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 XX
 PS Claim 25; Fig 11; 117pp; English.
 XX
 CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle and
 CC connecting peptide
 CC
 XX
 SQ Sequence 96 AA;
 QY
 DB 1 RRPWCYQV 9
 62 RRPWCYQV 70
 91.9%; Score 57; DB 5; Length 96;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Search completed: March 18, 2004, 11:16:23
 Job time : 4.13567 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:13:05 ; Search time 1.15974 Seconds
(without alignments)
445.152 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYQVQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep: *
6: /cgn2_6/prodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	91.9	89	4	US-09-101-272G-62
2	57	91.9	158	2	US-08-797-689-12
3	57	91.9	138	4	US-09-984-186-12
4	57	91.9	194	4	US-09-101-272G-80
5	57	91.9	200	4	US-09-101-272G-73
6	57	91.9	201	4	US-09-101-272G-96
7	57	91.9	208	4	US-09-101-272G-98
8	57	91.9	365	1	US-08-093-741-83
9	57	91.9	365	1	US-08-720-012-83
10	57	91.9	393	2	US-08-560-098A-44
11	57	91.9	393	3	US-08-967-024C-24
12	57	91.9	393	3	US-08-967-024C-25
13	57	91.9	411	1	US-08-087-163-1
14	57	91.9	411	1	US-08-286-748B-18
15	57	91.9	411	1	US-08-153-799-18
16	57	91.9	411	2	US-08-560-098A-48
17	57	91.9	411	3	US-09-181-816-1
18	57	91.9	411	4	US-09-403-736-2
19	57	91.9	430	1	US-07-942-157A-3
20	57	91.9	430	6	5219569-2
21	57	91.9	431	4	US-09-101-272G-1
22	57	91.9	431	6	5188829-1
23	57	91.9	432	2	US-08-560-098A-47
24	57	91.9	432	3	US-08-142-590B-25
25	57	91.9	432	3	US-08-148-910-12
26	57	91.9	432	3	US-08-148-910-12
27	57	91.9	432	3	US-08-148-910-12

28	40	64.5	83	2	US-08-811-949-2	Sequence 2, Appl
29	40	64.5	437	2	US-08-811-949-49	Sequence 49, Appl
30	40	64.5	437	2	US-08-811-949-51	Sequence 51, Appl
31	40	64.5	437	2	US-08-811-949-55	Sequence 55, Appl
32	40	64.5	437	2	US-08-811-949-57	Sequence 57, Appl
33	40	64.5	472	2	US-08-811-949-63	Sequence 63, Appl
34	40	64.5	477	2	US-08-560-098A-51	Sequence 51, Appl
35	40	64.5	527	1	US-07-609-510B-16	Sequence 16, Appl
36	40	64.5	527	2	US-08-811-949-39	Sequence 39, Appl
37	40	64.5	527	5	PCT-US91-01025A-2	Sequence 2, Appl
38	40	64.5	527	6	5185259-8	Sequence 2, Appl
39	40	64.5	527	6	5520913-1	Sequence 2, Appl
40	40	64.5	536	6	5200340-6	Sequence 40, Appl
41	40	64.5	562	2	US-08-811-949-43	Sequence 43, Appl
42	40	64.5	562	2	US-08-560-098A-50	Sequence 50, Appl
43	40	64.5	562	2	US-08-883-795A-38	Sequence 38, Appl
44	40	64.5	562	4	US-09-703-695A-4	Sequence 4, Appl
45	40	64.5	562	6	5185259-3	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-101-272G-62
Sequence 62, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Miasin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 62
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
US-09-101-272G-62

Query Match 91.9%; Score 57; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.0069;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQVQ 9
DB 67 RRPWCYQVQ 75

RESULT 2
US-08-797-689-12
Sequence 12, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleier, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guillon, Jean-Dominique
APPLICANT: Jung, Gerard
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Atcoia Road, 3C43
CITY: Colleegeville

STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-12

Query Match 91.9%; Score 57; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
|||||
Db 112 RRPWCYQV 120

RESULT 3
US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. 6686179
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 91.9%; Score 57; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
|||||
Db 112 RRPWCYQV 120

RESULT 4
US-09-101-272G-80
Sequence 80, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: 050979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 80
LENGTH: 194
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: ATRFI chimeric protein
US-09-101-272G-80

Query Match 91.9%; Score 57; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
|||||
Db 110 RRPWCYQV 118

RESULT 5
US-09-101-272G-73
Sequence 73, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR

FILE REFERENCE: 050979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.1
SEQ ID NO 73
LENGTH: 200
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: ATF domain of UPA
US-09-101-272G-73

Query Match 91.9%; Score 57; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 129 RRPWCYQV 137

RESULT 6
US-09-101-272G-96
Sequence 96, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: 050979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.1
SEQ ID NO 96
LENGTH: 201
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATPH1-CL chimeric protein
US-09-101-272G-96

Query Match 91.9%; Score 57; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 110 RRPWCYQV 118

RESULT 7
US-09-101-272G-98
Sequence 98, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: 050979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.1
SEQ ID NO 98
LENGTH: 208
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATPH1-ML chimeric protein
US-09-101-272G-98

Query Match 91.9%; Score 57; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 110 RRPWCYQV 118

RESULT 8
US-08-093-741-83
Sequence 83, Application US/08093741
Patent No. 5681721
GENERAL INFORMATION:
APPLICANT: STEFFENS, GERT J.
APPLICANT: WENDEL, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESSES:
ADDRESSER: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83

Query Match 91.9%; Score 57; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 63 RRPWCYQV 71

RESULT 9
US-08-720-012-83
Sequence 83, Application US/08720012
Patent No. 5747291

GENERAL INFORMATION:
APPLICANT: STEFFENS, GERD J.
APPLICANT: WRENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83
Query Match 91.9%; Score 57; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 63 RRPWCYQV 71
RESULT 10
US-08-560-098A-44
Sequence 44, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WRENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-44
Query Match 91.9%; Score 57; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 64 RRPWCYQV 72
RESULT 11
US-08-967-024C-24
Sequence 24, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WRENDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 91.9%; Score 57; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPCTVQV 9
DB 64 RRPCTVQV 72

RESULT 12
US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 613011
GENERAL INFORMATION:
APPLICANT: MENDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOGH, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKee, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 91.9%; Score 57; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPCTVQV 9
DB 64 RRPCTVQV 72

RESULT 13
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692

GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewich, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Faase, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04353/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
US-08-087-163-1

Query Match 91.9%; Score 57; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPCTVQV 9
DB 109 RRPCTVQV 117

RESULT 14
US-08-286-748B-18
Sequence 18, Application US/08286748B
Patent No. 5759542
GENERAL INFORMATION:
APPLICANT: Victor Gurewich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B

FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasae
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-18

Query Match 91.9%; Score 57; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 109 RRPWCYQV 117

RESULT 15

US-08-153-799-18

Sequence 18, Application US/08153799

Patent No. 5766883

GENERAL INFORMATION:

APPLICANT: Ballance, David J

APPLICANT: Goodey, Andrew R

TITLE OF INVENTION: Polypeptides

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: R Hain Swope, BOC Health Care Inc

STREET: 100 Mountain Avenue

CITY: Murray Hill

STATE: New Jersey

COUNTRY: USA

ZIP: 07974

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: Version #1.25

APPLICATION NUMBER: US/08/153,799

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/847975

FILING DATE: 06-MAR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 8909916.2

FILING DATE: 29-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB90/00650

FILING DATE: 26-APR-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/775952

FILING DATE: 29-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Swope, R Hain

REGISTRATION NUMBER: 24864

REFERENCE/DOCKET NUMBER: 92H832

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 665 2400

TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-799-18

Query Match 91.9%; Score 57; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 109 RRPWCYQV 117

Search completed: March 18, 2004, 11:21:23
Job time: 2.15974 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:19:25 ; Search time 2.97593 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYQVQ 10

Scoring table: BL0SUM62
Gap 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/PCRUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	10	13	US-10-076-421-5
2	57	91.9	86	15	US-10-233-675A-1
3	57	91.9	86	15	US-10-233-675A-1
4	57	91.9	86	15	US-10-233-675A-1
5	57	91.9	86	15	US-10-233-675A-1
6	57	91.9	86	15	US-10-233-675A-1
7	57	91.9	86	15	US-10-233-675A-1
8	57	91.9	86	15	US-10-233-675A-1
9	57	91.9	86	15	US-10-233-675A-1
10	57	91.9	86	15	US-10-233-675A-1
11	57	91.9	86	15	US-10-233-675A-1
12	57	91.9	86	15	US-10-233-675A-1
13	57	91.9	86	15	US-10-233-675A-1
14	57	91.9	86	15	US-10-233-675A-1
15	57	91.9	86	15	US-10-233-675A-1

16	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
17	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
18	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
19	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
20	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
21	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
22	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
23	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
24	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
25	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
26	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
27	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
28	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
29	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
30	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
31	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
32	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
33	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
34	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
35	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
36	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
37	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
38	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
39	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
40	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
41	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
42	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
43	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
44	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
45	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-10-076-421-5
Sequence 5, Application US/10076421
Publication No. US20020193304A1
GENERAL INFORMATION:
APPLICANT: WADA, MANABU
TITLE OF INVENTION: ANTI-HIV AGENTS
FILE REFERENCE: HAYAK-9
CURRENT APPLICATION NUMBER: US/10/076,421
PRIOR APPLICATION NUMBER: JP 2001-42655
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: JP 2001-184284
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 5
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapiens
US-10-076-421-5

Query Match 100.0%; Score 62; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 RRPWCYQVQ 10
US-10-233-675A-1
Sequence 1, Appl
Publication No. US20030228298A1
GENERAL INFORMATION:
APPLICANT: Neebit, Mark

```
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-1
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 3
US-10-233-675A-5
; Sequence 5, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-5
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 4
US-10-233-675A-7
; Sequence 7, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
```

```
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-7
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 5
US-10-233-675A-22
; Sequence 22, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human urokinase plasminogen activator
US-10-233-675A-22
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 6
US-10-233-675A-27
; Sequence 27, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
```

SOFTWARE: PatentIn version 3.1
SEQ ID NO 27
LENGTH: 86
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fragment of human urokinase plasminogen activator
NAME/KEY: MISC FEATURE
LOCATION: (43)-(43)
OTHER INFORMATION: Polymorphism - Xaa = Asn or Asp
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (74)-(74)
OTHER INFORMATION: Polymorphism - Xaa = Pro or Leu
US-10-233-675A-27

Query Match 91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 62 RRPWCYQV 70

RESULT 7
US-10-233-675A-9
Sequence 9, Application US/10233675A
Publication No. US20030228298A1
GENERAL INFORMATION:
APPLICANT: Neabitt, Mark
APPLICANT: Pong, Timothy
TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
FILE REFERENCE: ST01027
CURRENT APPLICATION NUMBER: US/10/233,675A
CURRENT FILING DATE: 2002-09-04
PRIOR APPLICATION NUMBER: 60/316,300
PRIOR FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 87
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: human derived fusion protein
US-10-233-675A-9

Query Match 91.9%; Score 57; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 63 RRPWCYQV 71

RESULT 8
US-10-233-675A-10
Sequence 10, Application US/10233675A
Publication No. US20030228298A1
GENERAL INFORMATION:
APPLICANT: Neabitt, Mark
APPLICANT: Pong, Timothy
TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
FILE REFERENCE: ST01027
CURRENT APPLICATION NUMBER: US/10/233,675A
CURRENT FILING DATE: 2002-09-04

PRIOR APPLICATION NUMBER: 60/316,300
PRIOR FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 87
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: human derived fusion protein
US-10-233-675A-10

Query Match 91.9%; Score 57; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 63 RRPWCYQV 71

RESULT 9
US-09-880-503-1
Sequence 1, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 88
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-1

Query Match 91.9%; Score 57; DB 9; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 62 RRPWCYQV 70

RESULT 10
US-09-880-503-9
Sequence 9, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 96
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-9

Query Match 91.9%; Score 57; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRPWCYQV 9
|||
Db 62 RRPWCYQV 70

RESULT 11
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: HIGHZL, Abd Al-RooF
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 91.9%; Score 57; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRPWCYQV 9
|||
Db 109 RRPWCYQV 117

RESULT 12
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL, BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689

FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: S792006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 91.9%; Score 57; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRPWCYQV 9
|||
Db 112 RRPWCYQV 120

RESULT 13
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US2003022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL, BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,667
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12

Query Match
Best Local Similarity 91.9%; Score 57; DB 14; Length 138;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 112 RRPWCYQV 120

RESULT 14
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match
Best Local Similarity 91.9%; Score 57; DB 14; Length 138;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 112 RRPWCYQV 120

RESULT 15
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match
Best Local Similarity 91.9%; Score 57; DB 14; Length 138;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYVQV 9
| | | | |
Db 112 RRPWCYVQV 120

Search completed: March 18, 2004, 11:36:07
Job time : 3.22593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 1.02845 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYVQVQ 10

Scoring table: BLOSUM62.
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	91.9	431	1 UKHU	u-plasminogen acti
2	57	91.9	433	1 UKBAY	u-plasminogen acti
3	57	91.9	442	1 UKPG	u-plasminogen acti
4	56	90.3	432	1 S18932	u-plasminogen acti
5	56	90.3	433	1 JN0560	u-plasminogen acti
6	53	85.5	433	1 UKMS	u-plasminogen acti
7	43	69.4	608	2 G82137	pyca protein VC194
8	43	69.4	655	1 A46688	hepatocyte growth
9	41	66.1	558	2 JCS878	plasma hyaluronan-
10	40	64.5	291	2 JCS878	conserved hypohet
11	40	64.5	304	2 B87639	u-plasminogen acti
12	40	64.5	394	2 JCS600	u-plasminogen acti
13	40	64.5	431	2 JCS059	u-plasminogen acti
14	40	64.5	434	1 A35005	u-plasminogen acti
15	40	64.5	477	1 A34369	u-plasminogen acti
16	40	64.5	477	2 JCS059	u-plasminogen acti
17	40	64.5	477	2 JCS059	u-plasminogen acti
18	40	64.5	559	1 A35029	u-plasminogen acti
19	40	64.5	559	1 A29941	u-plasminogen acti
20	40	64.5	560	1 JC4795	plasma hyaluronan-
21	40	64.5	562	1 UKHU	u-plasminogen acti
22	40	64.5	1113	1 A47106	myosin heavy chain
23	39	62.9	259	2 PN0687	cyclin - anthracno
24	39	62.9	411	2 I51285	hepatocyte growth
25	39	62.9	603	2 S28941	coagulation factor
26	39	62.9	615	1 KFHU12	coagulation factor
27	39	62.9	710	1 I51283	hepatocyte growth
28	39	62.9	711	1 A47136	macrophage-stimula
29	39	62.9	716	1 A40332	macrophage-stimula

30	39	62.9	716	1 JCS061	macrophage-stimula
31	39	62.9	728	1 JH0579	hepatocyte growth
32	39	62.9	728	1 A35644	hepatocyte growth
33	39	62.9	728	1 A60185	hepatocyte growth
34	39	62.9	806	2 T18840	hypothetical prote
35	39	62.9	946	1 A47239	hypothetical prote
36	39	62.9	1420	2 A32869	hypothetical prote
37	39	62.9	4548	1 S00657	hypothetical prote
38	38	61.3	89	2 A60140	plasma (EC 3.4.21
39	38	61.3	323	2 T25094	plasma (EC 3.4.21
40	38	61.3	593	2 S45281	hypothetical prote
41	37	59.7	157	2 T02034	coagulation factor
42	37	59.7	220	2 T46524	early light-induce
43	37	59.7	252	2 E75491	probable disulfide
44	37	59.7	409	2 B70636	probable fadB17 pr
45	37	59.7	656	2 S69554	hypothetical prote

ALIGNMENTS

RESULT 1
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen acti
in form
C;Species: Homo sapiens (man)
C;Date: 17-Dec-1982 #sequence, revision 04-Dec-1986 #text change 15-Sep-2000
C;Accession: A00931; I52209; J01012; A37561; I38102; S65783; A37562; A37563; A37564; A356
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasl, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A;Title: The human urokinase-plasminogen activator gene and its promoter.
A;Reference number: A00931; MUID:85215647; PMID:2987867
A;Accession: A00931
A;Molecule type: DNA
A;Residues: 1-431 <RC>
A;Cross-references: GB:K03427; NID:q340174; PIDN:CAA61257.1; PID:q340175
A;Note: the authors translated the codon ATG for residue 214 as Ile
R;Nagamine, Y.; Pearson, D.; Gratian, M.
Biochem. Biophys. Res. Commun. 135, 563-569, 1985
A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A;Reference number: I52209; MUID:86050639; PMID:3933505
A;Accession: I52209
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 145-161 <NAG1>
A;Cross-references: GB:K03427; NID:q340174; PIDN:AAA61257.1; PID:q340175
R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Aritamura, H.; Suyama, J
Gene 36, 183-188, 1985
A;Title: Molecular cloning of cDNA coding for human preprourokinase.
A;Reference number: J01012; MUID:86056954; PMID:2415429
A;Accession: J01012
A;Molecule type: mRNA
A;Residues: 1-213, 'I', 215-431 <NAG2>
A;Cross-references: GB:K03426; NID:q340155; PIDN:AA691138.1; PID:q340158; GB:J00244; NID
R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A;Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RN
A;Reference number: A37561; MUID:84272706; PMID:6589620
A;Accession: A37561
A;Molecule type: mRNA
A;Residues: 66-431 <VRP>
A;Cross-references: GB:J00244; NID:q3220138
R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen,
DNA 4, 139-146, 1985
A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prep
A;Reference number: I38102; MUID:85203359; PMID:3888571
A;Accession: I38102
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <UAC>
A;Cross-references: EMBL:X02760; NID:q35297; PIDN:CAA26535.1; PID:q35298

R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, K.
 A:Title: Biochem. Acta 1293, 83-89, 1996
 A:Title: Characterization of single chain urokinase-type plasminogen activator with a no
 A:Reference number: S65783; PMID:96186279; PMID:8652631
 A:Accession: S65783
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 21-140; 'L', 142-213; 'I', 215-431 <YOS>
 A:Cross-references: EMBL:D1143; NID:G131467; PIDN:BA01919.1; PID:G1199928
 R;Gunzler, W.A.; Steffens, G.J.; Oetting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A:Title: The primary structure of high molecular mass urokinase from human urine.
 A:Reference number: A37562; PMID:83055084; PMID:6754569
 A:Accession: A37562
 A:Molecule type: protein
 A:Residues: 21-177 <GUN>
 R;Schaller, J.; Nick, H.; Riekl, E.E.; Gilleesen, D.; Lergier, W.; Studer, R.O.
 Eur. J. Biochem. 125, 251-257, 1982
 A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and pre
 A:Reference number: A37563; PMID:83003608; PMID:6749491
 A:Accession: A37563
 A:Molecule type: protein
 A:Residues: 156-176; 179-193; 'T', 195; 'T', 197-224 <SCH>
 R;Steifens, G.J.; Gunzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A:Title: The complete amino acid sequence of low molecular mass urokinase from human uri
 A:Reference number: A37564; PMID:83055099; PMID:6754572
 A:Accession: A37564
 A:Molecule type: protein
 A:Residues: 158-410 <STE>
 R;Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A:Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinat
 A:Reference number: A35689; PMID:9036537; PMID:2393398
 A:Accession: A35689
 A:Molecule type: protein
 A:Residues: 21-30; 'X', 32, 'X', 34-38; 'X', 40-43 <KEN>
 A:Note: Identification of a fucose and attempt to determine its attachment site
 R;Rabhan, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazur, A.; Henkin, J.; Golitz
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
 A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
 A:Reference number: A36697; PMID:91097529; PMID:2125213
 A:Accession: A36697
 A:Molecule type: protein
 A:Residues: 21-34 <RAB>
 R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
 submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A51255; PDB:1KDU
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R;Li, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
 A:Reference number: A4375; PMID:93003110; PMID:1327118
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
 submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A66822; PDB:1URK
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A66058; PDB:1LWM
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C:Genetics:
 A:Gene: GDB:PLAU
 A:Cross-references: GDB:119497; OMTM:191840
 A:Map position: 10q24-10q24
 A:Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domin: signal sequence #status predicted <SIG>
 F:21-431/Product: urokinase-type plasminogen activator, single chain form #status predict
 F:21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MP
 F:31-62/Domin: EGF homology <EGF>
 F:70-151/Domin: kringle homology <KR>
 F:156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <MI
 F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental <MB
 F:179-419/Domin: trypsin homology <TRY>
 F:31-39; 33-51; 53-62; 70-151; 91-133; 122-146; 168-299; 209-225; 217-288; 313-382; 345-361; 372-400
 F:38/Binding site: carbohydrate (thr) (covalent) #status predicted
 F:178-179/Cleavage site: Lys-116 (plasmin) #status experimental
 F:224; 215; 376/Active site: His, Asp, Ser #status experimental
 F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 91.9%; Score 57; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.045;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
 Db 129 RRPWCYQV 137

RESULT 2
 UKBAY
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
 C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
 C:Accession: S14687; S08651
 R;Au, Y.P.T.; Wang, T.W.; Clowers, A.W.
 Nucleic Acids Res. 18, 3411, 1990
 A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
 A:Reference number: S14687; PMID:90287734; PMID:2113276
 A:Accession: S14687
 A:Molecule type: mRNA
 A:Residues: 1-433 <AUV>
 A:Cross-references: EMBL:X51935; NID:G38130; PIDN:CMA36200.1; PID:G38131
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domin: signal sequence #status predicted <SIG>
 F:21-176/Product: plasminogen activator chain A #status predicted <ACH>
 F:30-61/Domin: EGF homology <EGF>
 F:69-150/Domin: kringle homology <KR>
 F:178-433/Product: plasminogen activator chain B #status predicted <BCH>
 F:178-421/Domin: trypsin homology <TRY>
 F:167-298; 208-224; 216-287; 315-384; 347-363; 374-402/Disulfide bonds: #status predicted
 F:223; 274; 378/Active site: His, Asp, Ser #status predicted
 F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.9%; Score 57; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 0.046;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
 Db 128 RRPWCYQV 136

RESULT 3
 UKPG
 u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N:Alternate names: uPA
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R;Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; PMID:85087954; PMID:6096832
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240; 'H', 242-442 <NAG1>

A;Experimental source: kidney cell line LLC-PK1
R;Nagamine, Y.
Submitted to the Protein Sequence Database, December 1986
A;Reference number: A37566
A;Contents: annotation; correction to residue 241
C;Genetics:
A;Intons: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-198/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;190-430/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>
F;152/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;119-310,220-236,228-299,334-393,356-372,383-411/Disulfide bonds: #status predicted
F;235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 91.9%; Score 57; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 131 RRPWCYQV 139

RESULT 4
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N;Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
C;Accession: S24604; I60186; I53472; S18932
R;Rabban, S.A.
submitted to the EMBL Data Library, April 1992
A;Reference number: S24604
A;Accession: S24604
A;Molecule type: mRNA
A;Residues: 1-15, 'H', '17-23, 'G', '25-331, 'N', '333-432 <RAB>
A;Cross-references: EMBL:X65651; NID:957456; PIDN:CAA4601.1; PID:957457
A;Experimental source: tissue kidney
R;Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A;Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A;Reference number: I60186; MUID:92233409; PMID:1568219
A;Accession: I60186
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-432 <RES>
A;Cross-references: EMBL:X63434; NID:957465; PIDN:CAA5028.1; PID:957466
A;Experimental source: strain Fischer 344; tissue mammary
R;Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blas, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A;Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A;Reference number: I53472; MUID:92335549; PMID:1321734
A;Accession: I53472
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 31-62 <RE2>
A;Cross-references: EMBL:X66907; NID:9396200; PIDN:CAA47356.1; PID:9396279
C;Genetics:
A;Gene: uPA
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;119-332/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>
F;119-420/Domain: trypsin homology <TRY>
F;168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted

F;225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 90.3%; Score 56; DB 1; Length 432;
Best Local Similarity 88.9%; Pred. No. 0.067;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 129 RRPWCYQV 137

RESULT 5
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N;Alternate names: uPA
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: JN0560
R;Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A;Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
A;Reference number: JN0560; MUID:93216119; PMID:8385052
A;Accession: JN0560
A;Molecule type: mRNA
A;Residues: 1-433 <KRA>
A;Cross-references: GB:I03546; NID:9163800; PIDN:AAA51419.1; PID:9163801
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>
F;181-421/Domain: trypsin homology <TRY>
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 90.3%; Score 56; DB 1; Length 433;
Best Local Similarity 88.9%; Pred. No. 0.067;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 131 RRPWCYQV 139

RESULT 6
U005
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text_change 18-Jun-1999
C;Accession: A29420; A24615
R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A;Title: The murine urokinase-type plasminogen activator gene.
A;Reference number: A29420; MUID:88163489; PMID:2831940
A;Accession: A29420
A;Molecule type: DNA
A;Residues: 1-433 <DEG>
A;Cross-references: GB:M17922; NID:9202296; PIDN:AAA40539.1; PID:9202297
R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher,
Eur. J. Biochem. 148, 225-232, 1985
A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase-
A;Reference number: A24615; MUID:85179474; PMID:2985383
A;Accession: A24615
A;Molecule type: mRNA
A;Residues: 1-433 <BEL>
A;Cross-references: GB:I02389; NID:955127; PIDN:CAA26231.1; PID:955128
C;Genetics:
A;Intons: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F.1-20/Domain: signal sequence #status predicted <SIG>
F.32-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F.32-63/Domain: EGF homology <EGF>
F.71-152/Domain: kringle homology <KR>
F.180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>
F.180-421/Domain: trypsin homology <TRY>
F.169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F.126-277,378/Active site: His, Asp, Ser #status predicted

Query Match 85.5%; Score 53; DB 1; Length 433;
Best Local Similarity 77.8%; Pred. No. 0.21;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
:|||||:
Db 130 KRPWCYQV 138

RESULT 7
G82137
pvcA protein VC1949 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82137
R:Heideberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Drygol, I.; Sellers, H.
L.R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nucleotide 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82137
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-608 <HEI>
A:Cross-references: GB:AE004270; GB:AE003852; NID:g9656479; PIDN:AAFS097.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
A:Genetics:
A:Gene: VC1949
A:Map position: 1

Query Match 69.4%; Score 43; DB 2; Length 608;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPPWCYQV 9
:|||||:
Db 306 KPPWCYQV 313

RESULT 8
A46688
hepatocyte growth factor activator (BC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R:Miyaizawa, K.; Shimamura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII
A:Reference number: A46688; MUID:93252878; PMID:7683665
A:Accession: A46688
A:Molecule type: mRNA
A:Residues: 1-655 <MIY>
A:Cross-references: DBJ:U14012; NID:219680; PIDN:BA03113.1; PID:219681
A:Experimental source: liver (mRNA); serum (protein)
A:Note: sequence extracted from NCBI database (NCBI:131227, NCBI:131228)
A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c
C:Genetics:
A:Gene: GDB:HGFAC; HGFA; HGFAP
A:Cross-references: GDB:9954514
A:Map position: 4p16-4p16
C:Function:
A:Description: activates hepatocyte growth factor by specific proteolytic cleavage

A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F.1-34/Domain: signal sequence #status predicted <SIG>
F.108-148/Domain: fibronectin type II repeat homology <IF2>
F.164-157/Domain: EGF homology <EG1>
F.202-237/Domain: fibronectin type I repeat homology <IF1>
F.245-278/Domain: EGF homology <EG2>
F.286-367/Domain: kringle homology <KR>
F.373-407/Product: hepatocyte growth factor activator light chain #status experimental <L
F.408-555/Product: hepatocyte growth factor activator heavy chain #status experimental <H
F.406-641/Domain: trypsin homology <TRY>
F.40-48,290,468,492,546/Binding site: carbohydrate (Aan) (covalent) #status predicted
F.164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-362
F.447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 69.4%; Score 43; DB 1; Length 655;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPPWCYV 7
:|||||:
Db 346 RPPWCYV 351

RESULT 9
JC5878
plasma hyaluronan-binding protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C:Accession: JC5878
R:Hashimoto, K.; Tohe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.J
Biol. Pharm. Bull. 20, 1127-1130, 1997
A:Title: Cloning of the cDNA for a mouse homologue of human PHB: A novel hyaluronan-bin
A:Reference number: JC5878; MUID:98065239; PMID:9401717
A:Accession: JC5878
A:Molecule type: mRNA
A:Residues: 1-558 <HAS>
A:Comment: This protein acts as serine protease.
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsi
F.1-23/Domain: signal sequence #status predicted <SIG>
F.24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
F.75-106/Domain: EGF homology <EG1>
F.113-145/Domain: EGF homology <EG2>
F.152-185/Domain: EGF homology <EG3>
F.192-274/Domain: kringle homology <KR>
F.312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MAT
F.312-548/Domain: trypsin homology <TRY>

Query Match 66.1%; Score 41; DB 2; Length 558;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPPWCYQV 9
:|||||:
Db 252 KPPWCYQV 259

RESULT 10
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PluAT) cDNA obtained from human endo.
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DDBT
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>

c;superfamily: class plasminogen activator; EGF homology; fibronectin type 1 repeat homology

A/Molecule type: mRNA

A;Residues: 1-434 <LES>

A;Cross-references: GB:J05187, NID:g212858, PIDN:AAA9131.1, PID:g212859
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F;40-71/Domain: EGF homology <EGF>
 F;79-158/Domain: Kringle homology <KR>
 F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F;173-416/Domain: trypsin homology <TRY>
 F;162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
 F;217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 64.5%; Score 40; DB 1; Length 434;
 Best Local Similarity 71.4%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RPMCYVQ 8
 : ||||| :
 Db 136 RPMCYTK 142

RESULT 15

A34369
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
 C;Species: Megaderma lyra
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: A34369
 R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
 J. Biol. Chem. 264, 17947-17952, 1989
 A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
 A;Reference number: A34369; MUID:90036867; PMID:2509450
 A;Accession: A34369
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-477 <GAR>
 A;Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator #status predicted <PLA>
 F;42-79/Domain: fibronectin type I repeat homology <IFA>
 F;87-120/Domain: EGF homology <EGF>
 F;128-209/Domain: kringle homology <KR>
 F;226-477/Domain: trypsin homology <TRY>
 F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
 F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 64.5%; Score 40; DB 1; Length 477;
 Best Local Similarity 83.3%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPMCYV 7
 : ||||| :
 Db 188 RPMCYV 193

Search completed: March 18, 2004, 11:20:19
 Job time : 2.02845 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:54:45 ; Search time 0.612691 Seconds
(without alignments)
849.859 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPACTYQVQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	57	91.9	431 1	UROC_HUMAN
2	57	91.9	433 1	UROC_PAPCY
3	57	91.9	442 1	UROC_PIG
4	56	90.3	432 1	UROC_RAT
5	56	90.3	433 1	UROC_BOVIN
6	53	85.5	433 1	UROC_MOUSE
7	43	69.4	653 1	HGFA_MOUSE
8	43	69.4	655 1	HGFA_HUMAN
9	40	64.5	394 1	URTB_DESRO
10	40	64.5	431 1	URTB_DESRO
11	40	64.5	434 1	UROC_CHICK
12	40	64.5	477 1	URTB_DESRO
13	40	64.5	477 1	URTB_DESRO
14	40	64.5	559 1	TPA_MOUSE
15	40	64.5	559 1	TPA_RAT
16	40	64.5	562 1	TPA_HUMAN
17	40	64.5	566 1	TPA_BOVIN
18	40	64.5	1109 1	MYSD_DICDI
19	39	62.9	245 1	CGI_COGLI
20	39	62.9	452 1	KRM1_XENLA
21	39	62.9	461 1	KRM2_MOUSE
22	39	62.9	462 1	KRM2_HUMAN
23	39	62.9	603 1	FA12_CAVPO
24	39	62.9	615 1	FA12_HUMAN
25	39	62.9	711 1	HGFL_HUMAN
26	39	62.9	716 1	HGFL_MOUSE
27	39	62.9	728 1	HGF_HUMAN
28	39	62.9	728 1	HGF_MOUSE
29	39	62.9	728 1	HGF_RAT
30	39	62.9	1420 1	APOA_MACMU
31	39	62.9	4548 1	APOA_HUMAN
32	38	61.3	473 1	KRM1_MOUSE
33	38	61.3	473 1	KRM1_RAT

34	38	61.3	475 1	KRM1_HUMAN	Q6m88 homo sapien
35	38	61.3	593 1	FA12_BOVIN	P8140 bos taurus
36	37.5	60.5	862 1	TASY_TAXBA	Q93ya3 taxus baccia
37	37.5	60.5	862 1	TASY_TAXBR	Q41594 taxus brevis
38	37.5	60.5	862 1	TASY_TAXCH	Q95c37 taxus chine
39	37	59.7	416 1	ORF5_STRGR	Q54228 streptomyc
40	37	59.7	416 1	Y516_STRCO	Q94414 streptomyc
41	37	59.7	512 1	BETC_RHIME	Q69787 rhizobium m
42	37	59.7	656 1	PUP6_YEAST	Q04373 saccharomyc
43	36	58.1	333 1	PLMN_CANPA	P80009 canis famli
44	36	58.1	374 1	SPOP_HUMAN	Q43791 homo sapien
45	36	58.1	758 1	CC27_YEAST	P38042 saccharomyc

ALIGNMENTS

RESULT 1
UROC_HUMAN STANDARD; PRT; 431 AA.
ID UROC_HUMAN AC P00749; Q15844; Q1618; Q969W6;
DC 21-JUN-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RT Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia
RT coli.";
RN Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RT Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
RT Gene 36:183-186(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=388571;
RA Jacobs P., Cravador A., Lortien R., Brockly F., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of
RT human prepro-urokinase cDNA.";
RT DNA 4:135-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=2388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toehly S., Carrinci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehy J., Helton A., Kettaman C., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [17]
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.,
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8]
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.,
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller R.O., Nick H., Rickli E.E., Gilleesen D., Lergier W.,
 RA Studer R.O.,
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.,
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.,
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Speil S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.,
 RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.,
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.,
 RT "The structure of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).

RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=9419701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.,
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Yamaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hanada K.,
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Comme B., Berczy M., Belin D.,
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Comme B., Berczy M., Belin D.,
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.,
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: Potent plasminogen activator and is clinically used for
 CC therapy of thrombolytic disorders.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain.
 CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X02419; CAA26868.1; -;
 CC EMBL, M15476; AAA61253.1; -;
 CC EMBL, D00244; BAA00175.1; -;
 CC EMBL, D11443; BAA01919.1; -;
 CC EMBL, X02760; CAA26535.1; -;
 CC EMBL, AF377330; AAK53822.1; -;
 CC EMBL, BC033575; AAH13575.1; -;
 CC EMBL, K03226; AAC97138.1; -;
 CC EMBL, K02286; AAA61252.1; -;
 CC EMBL, A21571; CAA01559.1; -;
 CC EMBL, A18397; CAA01390.1; -;
 CC PIR, A00931; UKHU.
 CC PDB, 1KDU; 31-OCT-93.

Query Match 91.9%; Score 57; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPCTYOV 9
 Db 129 RRPCTYOV 137

RESULT 2
 ID UROK_PAPCY STANDARD; PRT; 433 AA.
 AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Papio.
 OC NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA Au Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RT plasminogen activator."
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain (by similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activator; Hydrolyase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1
 FT CHAIN 21 433
 FT CHAIN 21 176
 FT CHAIN 155 176
 FT CHAIN 178 433
 FT DOMAIN 26 62
 FT DOMAIN 69 150
 FT DOMAIN 151 177
 FT DOMAIN 178 433
 FT DISULFID 30 38
 FT DISULFID 32 50
 FT DISULFID 52 61
 FT DISULFID 167 298
 FT DISULFID 208 224
 FT DISULFID 216 287
 FT DISULFID 315 384
 FT DISULFID 347 363
 FT ACT_SITE 223 223
 FT ACT_SITE 274 274
 FT ACT_SITE 378 378
 FT CARBOHYD 324 324
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFDCC8792 CRC64; (BY SIMILARITY).

Query Match 91.9%; Score 57; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPCTYOV 9
 Db 128 RRPCTYOV 136

RESULT 3
 ID UROK_PIG STANDARD; PRT; 442 AA.
 AC P04185;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=85087954; PubMed=6096832;
 RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator."
 RL Nucleic Acids Res. 12:9525-9541(1984).
 RN [2]
 RP REVISION TO 241.
 RA Nagamine Y.;
 RL Submitted (DEC-1986) to the PIR data bank.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).

 DR EMBL; X01648; CAA25806.1; --
 DR EMBL; X02724; CAA26511.1; --
 DR PIR; A00932; UKPG.
 DR HSSP; P00749; IKDU.
 DR MEROPS; S01.231; --
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; UKR_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRODOM; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 1 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
 FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 189 CONNECTING PEPTIDE.
 FT DOMAIN 190 442 SERINE PROTEASE.
 FT CAROHD 152 152 N-LINKED (GLCNAC. . .).
 FT DISULFD 33 41 BY SIMILARITY.
 FT DISULFD 35 53 BY SIMILARITY.
 FT DISULFD 55 64 BY SIMILARITY.
 FT DISULFD 179 310 INTERCHAIN (BY SIMILARITY).
 FT DISULFD 220 236 BY SIMILARITY.
 FT DISULFD 228 299 BY SIMILARITY.
 FT DISULFD 324 393 BY SIMILARITY.
 FT DISULFD 356 372 BY SIMILARITY.
 FT DISULFD 363 411 BY SIMILARITY.
 FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
 FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
 FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
 FT ACT_SITE 241 241 Q -> H (IN REF. 1; CAA25806).
 FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
 FT CONFLICT 288 288 E32PCEFS0131EE CRC64;
 SQ SEQUENCE 442 AA; 49116 MW; 49116 MW;

Query Match 91.9%; Score 57; DB 1; Length 442;
 Best Local Similarity 100.0%; Pred. No. 0.015; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

OY 1 RRPWCYGV 9
 |||||
 DB 131 RRPWCYGV 139

RESULT 4
 UROK_RAT

ID UROK_RAT STANDARD; PRT; 432 AA.
 AC P29598;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer 344;
 RX MEDLINE=92233409; PubMed=1568219;
 RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
 RA Kellford R.F.;
 RT "Transcriptional and posttranscriptional activation of urokinase
 RT plasminogen activator gene expression in metastatic tumor cells";
 RL Cancer Res. 52:2489-2496(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (by similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).

 DR EMBL; X63434; CAA45028.1; --
 DR EMBL; X65651; CAA46601.1; --
 DR PIR; S24604; S18932.
 DR HSSP; P00749; IKDU.
 DR MEROPS; S01.231; --
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; UKR_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KM Kringling; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
 FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
 FT DOMAIN 27 63 EGF-LIKE.
 FT DOMAIN 152 178 KRINGLE.
 FT DOMAIN 179 432 CONNECTING PEPTIDE.
 FT DISULFID 31 39 SERINE PROTEASE.
 FT DISULFID 33 51 BY SIMILARITY.
 FT DISULFID 53 62 BY SIMILARITY.
 FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 210 226 BY SIMILARITY.
 FT DISULFID 218 289 BY SIMILARITY.
 FT DISULFID 314 383 BY SIMILARITY.
 FT DISULFID 346 362 BY SIMILARITY.
 FT DISULFID 373 401 BY SIMILARITY.
 FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
 FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
 FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
 FT CONFLICT 16 16 N -> H (IN REF. 2).
 FT CONFLICT 24 24 E -> G (IN REF. 2).
 FT CONFLICT 332 332 D -> N (IN REF. 2).
 SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 90.3%; Score 56; DB 1; Length 432;
 Best Local Similarity 88.9%; Pred. No. 0.021;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
 Db 129 RRPWCYQI 137

RESULT 5
 UROK_BOVIN STANDARD; PRT; 433 AA.
 ID UROK_BOVIN 005589; 028209;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Arctic endothelium;
 RX MEDLINE=93216119; PubMed=8385052;
 RA Kaetzelmar J., Haendler B., Kojima S., Rifkin D.B.,
 RA Schlaun W.-D.,
 RT "Bovine urokinase-type plasminogen activator and its receptor:
 RT cloning and induction by retinoic acid.";
 RL Gene 125:177-183(1993).
 RN [2]
 RP SEQUENCE OF 12-433 FROM N.A.
 RC TISSUE=Kidney;
 RA Ravn P., Berglund L., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen activators uPA
 RT and tPA.";
 RL Int. Dairy J. 5:605-617(1995).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- INDUCTION: By retinoic acid.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.

CC -1- SIMILARITY: Contains 1 kringling domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; L03546; AAA51419.1; -;
 DR EMBL; X85801; CAAS9796.1; -;
 DR EMBL; J05560; J05560.
 DR HSP; P00749; 1LMW.
 DR MEROPS; S01.231; -;
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000001; Kringling.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR01254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; Kringling_1.
 DR Pfam; PF00088; trypsin_1.
 DR PIRSF; PIRSF00114; Urokinase_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringling_1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS0186; EGF_2; FALSE_NEG.
 DR PROSITE; PS0026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KM Kringling; EGF-like domain; Signal; Zymogen.
 FT SIGNAL 1 20
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
 FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 180 CONNECTING PEPTIDE.
 FT DOMAIN 181 433 SERINE PROTEASE.
 FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 290 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 189 189 A -> T (IN REF. 2).
 SQ SEQUENCE 433 AA; 48730 MW; 4DE1B9D4DA47027A CRC64;

Query Match 90.3%; Score 56; DB 1; Length 433;
 Best Local Similarity 88.9%; Pred. No. 0.022;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
 Db 131 RRPWCYQI 139

RESULT 6
 UROK_MOUSE

ID	PROK_MOUSE	STANDARD;	ERR;	433 AA.
AC	P06863;			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DT	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator).			
GN	PLAU.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=851179474; PubMed=2985383;			
RA	Belin D., Vassalli J.-D., Compehne C., Godeau F., Ngamane Y.,			
RA	Reich E., Koehler H.P., Duvolsin R.M.;			
RT	"Cloning, nucleotide sequencing and expression of cDNAs encoding			
RT	mouse urokinase-type plasminogen activator.";			
RL	Eur. J. Biochem. 148:225-232(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86163489; PubMed=2831940;			
RA	Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;			
RT	"The murine urokinase-type plasminogen activator gene.";			
RL	Biochemistry 26:8270-8279(1987).			
CC	-1- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in			
CC	plasminogen to form plasmin.			
CC	-1- SUBUNIT: Found in high and low molecular mass forms. Each consists			
CC	of two chains, A and B. The high molecular mass form contains a			
CC	long chain A. Cleavage occurs after residue 156 in the low			
CC	molecular mass form to yield a short A1 chain (By similarity).			
CC	-1- SIMILARITY: Belongs to peptidase family S1.			
CC	-1- SIMILARITY: Contains 1 EGF-like domain.			
CC	-1- SIMILARITY: Contains 1 kringle domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X02389; CAA26231.1; -			
DR	EMBL; M17922; AAA40539.1; -			
DR	PIR; A29420; UKMS.			
DR	HSSP; P00749; IKDU.			
DR	MEROPS; S01.231; -.			
DR	MGD; MGI:97611; Plau.			
DR	InterPro; IPR009003; Cys_Ser_trypsin.			
DR	InterPro; IPR006209; EGF_Like.			
DR	InterPro; IPR006210; IEGF.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR008293; Pept_S1A_uPA.			
DR	InterPro; IPR001254; Peptidase_S1.			
DR	InterPro; IPR001314; Peptidase_S1A.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PIRSF; PIRSF001144; Urk_plasm_act; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000395; KRINGLE; 1.			
DR	SMART; SM00181; EGF; 1.			
DR	SMART; SM00130; KR; 1.			
DR	SMART; SM00020; TYP_Spc; 1.			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS01186; EGF_2; FALSE_NEG.			
DR	PROSITE; PS50026; EGF_3; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS50070; KRINGLE_2; 1.			
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.			

DR	PROSITE, PS00135 : TRYPSIN_SER_1.
KW	Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
XN	Kringle; EGF-like domain; Zymogen; Signal.
FT	SIGNAL
FT	CHAIN
FT	CHAIN
FT	CHAIN
FT	CHAIN
FT	CHAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	ACT SITE
FT	ACT SITE
FT	ACT_SITE
SC	SEQUENCE
QY	1 RRPWCYGV 9
DB	130 KRPWCYGV 138
Matches	7; Conservative 2; Mismatches 0; Indels 0; Gaps
Score	85.5%; DB 1; Length 433;
Pred.	77.8%; No. 0.069;
Result	7
HGFA_MOUSE	STANDARD; PRT; 653 AA.
ID	HGFA_MOUSE
AC	Q9R098; Q9JXV4;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
GN	HGFAC.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_Taxid=10090;
RN	(1)
RP	SEQUENCE FROM N.A.
RC	STRAIN=BALB/c;
RA	Itoh H., Kataoka H., Koono H.;
RT	"Mouse hepatocyte growth factor activator."
RL	Submitted (MAR-1993) to the EMBL/Genbank/DBDJ databases. (2)
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21226753; PubMed=11032833;
RA	van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang Y., Huan Y.;
RT	"Activation of hepatocyte growth factor (HGF) by endogenous HGF activator is required for metanephric kidney morphogenesis in vitro."
RL	J. Biol. Chem. 276:15099-15106(2001).
CC	- FUNCTION: Activates hepatocyte growth factor (HGF) by converting it from a single chain to a heterodimeric form (By similarity).
CC	- SUBUNIT: Dimer of a short chain and a long chain linked by a disulfide bond (By similarity).
CC	- SUBCELLULAR LOCATION: Secreted as an inactive single-chain precursor and is then activated to a heterodimeric form (By similarity).
CC	- SIMILARITY: Belongs to peptidase family S1.
CC	- SIMILARITY: Contains 2 EGF-like domains.

CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

CC -----
 CC EMBL: AF09017; AAF02489.1; -
 CC EMBL: AF224724; AAF34712.1; -
 CC HSSP: P00763; IDPO.
 CC MGD: MGI:1859281; Hgfac.
 CC DR INTERPRO: IPR003003; Cys_Ser_tyrpsin.
 CC DR INTERPRO: IPR000742; EGF_2.
 CC DR INTERPRO: IPR006209; EGF_1like.
 CC DR INTERPRO: IPR000083; Fibronectin.
 CC DR INTERPRO: IPR000562; FN_Type_II.
 CC DR INTERPRO: IPR006210; IEGF.
 CC DR INTERPRO: IPR000001; Kringle.
 CC DR INTERPRO: IPR001254; Peptidase_S1.
 CC DR INTERPRO: IPR001314; Peptidase_S1A.
 CC DR Pfam: PF00008; EGF_2.
 CC DR Pfam: PF00039; fn1; 1.
 CC DR Pfam: PF00040; fn2; 1.
 CC DR Pfam: PF00051; kringle; 1.
 CC DR Pfam: PF00089; trypsin; 1.
 CC DR PRINTS: PR00722; CHYMOTRYPSIN.
 CC DR PRINTS: PR00013; FNTYPEII.
 CC DR PRINTS: PR00018; KRINGLE.
 CC DR PRODOM: PD000995; FN_Type_II; 1.
 CC DR PRODOM: PD000395; Kringle; 1.
 CC DR SMART: SM00181; EGF_2.
 CC DR SMART: SM00059; FN2; 1.
 CC DR SMART: SM00130; KR; 1.
 CC DR SMART: SM00020; TYP_SPC; 1.
 CC DR PROSITE: PS00022; EGF_1; 2.
 CC DR PROSITE: PS01186; EGF_2; 1.
 CC DR PROSITE: PS50026; EGF_3; 2.
 CC DR PROSITE: PS01253; FIBRONECTIN_1; 1.
 CC DR PROSITE: PS00023; FIBRONECTIN_2; 1.
 CC DR PROSITE: PS00021; KRINGLE_1; 1.
 CC DR PROSITE: PS50070; KRINGLE_2; 1.
 CC DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 CC DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
 CC KM Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 CC EGF-like domain; Repeat; Zymogen.
 CC FT SIGNAL 1 29
 CC FT PROPEL 30 369
 CC FT CHAIN 370 405
 CC FT CHAIN 406 653
 CC FT DOMAIN 105 145
 CC FT DOMAIN 157 195
 CC FT DOMAIN 197 237
 CC FT DOMAIN 238 276
 CC FT DOMAIN 283 364
 CC FT DOMAIN 406 653
 CC FT ACT_SITE 445 445
 CC FT ACT_SITE 495 495
 CC FT ACT_SITE 596 596
 CC FT DISULFID 105 130
 CC FT DISULFID 119 145
 CC FT DISULFID 161 172
 CC FT DISULFID 166 183
 CC FT DISULFID 185 194
 CC FT DISULFID 199 227

FT DISULFID 225 234 BY SIMILARITY.
 FT DISULFID 242 253 BY SIMILARITY.
 FT DISULFID 247 264 BY SIMILARITY.
 FT DISULFID 266 275 BY SIMILARITY.
 FT DISULFID 283 364 BY SIMILARITY.
 FT DISULFID 304 346 BY SIMILARITY.
 FT DISULFID 335 359 BY SIMILARITY.
 FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 430 446 BY SIMILARITY.
 FT DISULFID 438 508 BY SIMILARITY.
 FT DISULFID 533 602 BY SIMILARITY.
 FT DISULFID 565 581 BY SIMILARITY.
 FT DISULFID 592 620 BY SIMILARITY.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 164 164 G -> W (IN REF. 2).
 SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

Query Match 69.4%; Score 43; DB 1; Length 653;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPMCYV 7
 Db 343 RPMCYV 348

RESULT 8
 HGFA HUMAN STANDARD; PRT; 655 AA.
 ID HGFA HUMAN
 AC Q04756; Q14726;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hepatocyte growth factor activator precursor (BC 3.4.21.-) (HGF
 DE activator) (HGFA).
 GN HGFA.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum;
 RX MEDLINE=93252878; PubMed=768365;
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.;
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor. Structural similarity for the protease precursor to blood
 RT coagulation factor XII.";
 RL J. Biol. Chem. 268:10024-10028 (1993).
 RN [2]
 RP SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odeh C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Activates hepatocyte growth factor (HGF) by
 CC converting it from a single chain to a heterodimeric form.
 CC -1- SUBUNIT: Dimer of a short chain and a long chain linked by a
 CC disulfide bond.
 CC -1- SUBCELLULAR LOCATION: Secreted as an inactive single-chain
 CC precursor and is then activated to a heterodimeric form.
 CC -1- TISSUE SPECIFICITY: Liver.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -1- CAUTION: It is uncertain whether Met-1 is the initiator.


```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63990; AAA31595.1; -.
CC PIR; J50600; J50600.
CC HSSP; P98119; 1A51.
CC MEROPS; S01.239; -.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR01254; Peptidase_S1.
CC InterPro; IPR01314; Peptidase_S1A.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPsin_DOM; 1.
CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC PROSITE; PS00135; TRYPsin_SER; 1.
CC K1M; Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
CC Kringle; Signal; Multigene family.
CC FT SIGNAL 1 36 POTENTIAL.
CC FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
CC FT DOMAIN 45 126 KRINGLE.
CC FT ACT_SITE 142 394 SERINE PROTEASE.
CC FT ACT_SITE 189 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 45 126 BY SIMILARITY.
CC FT DISULFID 66 108 BY SIMILARITY.
CC FT DISULFID 97 121 BY SIMILARITY.
CC FT DISULFID 131 262 BY SIMILARITY.
CC FT DISULFID 174 190 BY SIMILARITY.
CC FT DISULFID 182 251 BY SIMILARITY.
CC FT DISULFID 276 351 BY SIMILARITY.
CC FT DISULFID 308 324 BY SIMILARITY.
CC FT DISULFID 341 369 BY SIMILARITY.
CC FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3DB1FCD CRC64;
CC -----
CC Query Match 64.5%; Score 40; DB 1; Length 394;
CC Best Local Similarity 83.3%; Pred. No. 10;
CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 2 RFWCYV 7
CC Db 105 KPMCYV 110
CC -----
CC RESULT 10
CC URTB_DESCRO STANDARD; PRT; 431 AA.
CC ID URTB_DESCRO
CC AC P98121;
CC DT 01-FEB-1996 (Rel. 33, Created)
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
CC beta).
CC OS Desmodus rotundus (Vampire bat).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
CC OC Desmodontinae; Desmodus.
CC NCBI_taxid=9430;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC ISSUES=Salivary gland;
RX MEDLINE=2039036; PubMed=1937019;
RA Kretschmar J., Haendler B., Langer G., Boïdol W., Bringmann P.,
RA Alagon A., Donner P., Schlenker W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237 (1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=9339059; PubMed=1309059;
RA Schlenker W.-D., Alagon A., Boïdol W., Bringmann P., Petri T.,
RA Kretschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403 (1992)
CC -1 FUNCTION: Probably essential to support the feeding habits of this
CC exclusively haematophagous animal. Probable potent thrombolytic
CC agent.
CC -1 CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1 SUBUNIT: Monomer.
CC -1 SIMILARITY: Belongs to peptidase family S1.
CC -1 SIMILARITY: Contains 1 EGF-like domain.
CC -1 SIMILARITY: Contains 1 Kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63989; AAA31594.1; -.
CC PIR; J50599; J50599.
CC HSSP; P98119; 1A51.
CC MEROPS; S01.239; -.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR00134; Peptidase_S1A.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS50026; EGF_3; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPsin_DOM; 1.
CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC PROSITE; PS00135; TRYPsin_SER; 1.
CC K1M; Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Signal; Multigene family.
CC FT SIGNAL 1 36 POTENTIAL.
CC FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
CC FT DOMAIN 82 163 KRINGLE.
CC FT DOMAIN 179 431 SERINE PROTEASE.
CC FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).

```

```

FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.
FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 431 AA; 48321 MW; 699B5E675B162CBF CRC64;

Query Match 64.5%; Score 40; DB 1; Length 431;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPMCIV 7
Db 142 RPMCIV 147

RESULT 11
UROK_CHICK STANDARD; PRT; 434 AA.
ID UROK_CHICK
AC P15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
OS (u-plasminogen activator).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90110185; PubMed=2295632;
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RT "The chicken urokinase-type plasminogen activator gene.";
RL J. Biol. Chem. 265:1339-1344(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC
CC EMBL; J05187; AAA49131.1; -.
CC EMBL; J05188; AAA49130.1; -.
CC PIR; A35005; A35005.
CC HSSP; P00763; IDRO.
CC MEROPS; S01.231; -.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR008293; Pept_S1A_uPA.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.

```

```

DR PIRSF; PIRSF001144; Urokinase_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TYP_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS0186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).
FT DOMAIN 36 72 EGF-LIKE.
FT DOMAIN 79 158 KRINGLE.
FT DOMAIN 159 172 CONNECTING PEPTIDE.
FT DOMAIN 173 434 SERINE PROTEASE.
FT DISULFID 40 48 BY SIMILARITY.
FT DISULFID 42 60 BY SIMILARITY.
FT DISULFID 62 71 INTERCHAIN (BY SIMILARITY).
FT DISULFID 162 296 BY SIMILARITY.
FT DISULFID 202 218 BY SIMILARITY.
FT DISULFID 210 285 BY SIMILARITY.
FT DISULFID 310 379 BY SIMILARITY.
FT DISULFID 342 358 BY SIMILARITY.
FT DISULFID 369 397 BY SIMILARITY.
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 434 AA; 49400 MW; BD881048D666A55 CRC64;

Query Match 64.5%; Score 40; DB 1; Length 434;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RPMCIV 8
Db 136 RPMCIV 142

RESULT 12
URTI_DESRO STANDARD; PRT; 477 AA.
ID URTI_DESRO
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kretschnar J., Haendler B., Langer G., Boidl W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]

```

RP CHARACTERIZATION.
 RX MEDLINE=9339059; PubMed=1309059;
 RA Schlemming W.-D., Alagon A., Boidon W., Bringmann P., Petri T.,
 RA Kaerzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 RN (3)
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RC TISSUE=Salivary gland;
 RX MEDLINE=9602274; PubMed=9354616;
 RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
 RA Schlemming W.D., Bode W.;
 RT "Catalytic domain structure of vampire bat plasminogen activator: a
 RT molecular paradigm for proteolysis without activation cleavage.";
 RL Biochemistry 36:13483-13493(1997).
 CC -1- FUNCTION: Probably essential to support the feeding habits of this
 CC exclusively haematophagous animal. Potent thrombolytic agent.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
 CC plasminogen to form plasmin.
 CC -1- ENZYME REGULATION: Activity toward plasminogen is stimulated in
 CC the presence of fibrin I.
 CC -1- SUBUNIT: Monomer.
 CC -1- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
 CC and the kringle domain apparently mediates fibrin-induced
 CC stimulation of activity.
 CC -1- SIMILARITY: Belongs to EGF-like domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: M63987; AAA31591.1; -;
 DR EMBL: M63986; AAA31592.1; -;
 DR PIR: J50597; J50597.
 DR PDB: 1A5I; 23-MAR-99.
 DR MEROPS: S01.232; -;
 DR GlycoSuiteDB: P98119; -;
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR006083; Fibrinctn1.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00039; fn1; 1.
 DR Pfam: PF00051; Kringle_1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00016; KRINGLE.
 DR ProDom: PD000395; Kringle_1.
 DR SMART: SM00181; EGF_1.
 DR SMART: SM00058; FN1; 1.
 DR SMART: SM00130; KR_1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS50026; EGF_3; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 36
 FT CHAIN 37 477
 FT DOMAIN 40 82
 FT DOMAIN 83 121
 FT DOMAIN 128 209
 FT DOMAIN 225 477
 FT ACT_SITE 272 272
 FT ACT_SITE 321 321
 FT ACT_SITE 428 428
 FT DISULFID 42 72
 FT DISULFID 70 79
 FT DISULFID 87 98
 FT DISULFID 92 109
 FT DISULFID 111 120
 FT DISULFID 128 209
 FT DISULFID 149 191
 FT DISULFID 180 204
 FT DISULFID 214 345
 FT DISULFID 257 273
 FT DISULFID 265 334
 FT DISULFID 359 434
 FT DISULFID 391 407
 FT DISULFID 424 452
 FT CARBOHYD 153 153
 FT CARBOHYD 398 398
 FT TRN 214 215
 FT STRAND 223 224
 FT TURN 226 227
 FT STRAND 230 231
 FT HELIX 234 236
 FT TURN 238 239
 FT STRAND 240 245
 FT STRAND 254 263
 FT TURN 264 265
 FT STRAND 266 269
 FT HELIX 271 273
 FT TURN 280 282
 FT STRAND 284 287
 FT TURN 297 298
 FT STRAND 300 309
 FT TURN 311 312
 FT TURN 315 317
 FT TURN 319 320
 FT STRAND 323 328
 FT STRAND 338 338
 FT TURN 339 340
 FT STRAND 341 341
 FT STRAND 345 345
 FT TURN 349 350
 FT TURN 355 356
 FT STRAND 358 363
 FT STRAND 366 366
 FT STRAND 374 374
 FT STRAND 379 385
 FT HELIX 388 390
 FT TURN 393 398
 FT TURN 403 404
 FT STRAND 405 409
 FT TURN 425 426
 FT TURN 428 429
 FT STRAND 431 436
 FT TURN 437 438
 FT STRAND 439 448
 FT TURN 455 456
 FT STRAND 459 463
 FT HELIX 464 467
 FT HELIX 468 474
 SQ SEQUENCE 477 AA; 53616 MW; AA06FDI739C10B5E CRC64;
 N-LINKED (GLCNAC. . .)
 /FTId=CAR_000027.
 N-LINKED (GLCNAC. . .)
 /FTId=CAR_000028.

```

Query Match      64.5%; Score 40; DB 1; Length 477;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy      2 RPKCV 7
        :|||||
Db      188 KPMCTV 193

RESULT 13
UR12 DESRO      STANDARD; PRT; 477 AA.
AC  p15638;
DT  01-APR-1990 (Rel. 14, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
DE  alpha-2) (BAT-PA) (T-plasminogen activator).
OS  Desmodus rotundus (Vampire bat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC  Desmodontinae; Desmodus.
OX  NCBI_TaxID=9430;
RN  [1]
RP  TISSUE=Salivary gland;
RC  MEDLINE=92039036; PubMed=1937019;
RA  Krietzschmar J., Haendler B., Langer G., Boisdol W., Bringmann P.,
RA  Alagon A., Donner P., Schleuning W.D.;
RT  "The plasminogen activator family from the salivary gland of the
RT  vampire bat Desmodus rotundus: cloning and expression.";
RL  Gene 105:229-237(1991).
RN  [2]
RP  TISSUE=Salivary gland;
RC  MEDLINE=90036867; PubMed=2509450;
RA  Gadel S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA  Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT  "Isolation, characterization, and cDNA cloning of a vampire bat
RT  salivary plasminogen activator.";
RL  J. Biol. Chem. 264:17947-17952(1989).
RN  [3]
RP  CHARACTERIZATION.
RC  MEDLINE=9339059; PubMed=1309059;
RA  Schleuning W.D., Alagon A., Boisdol W., Bringmann P., Petri T.,
RA  Krietzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
RA  Donner P.;
RT  "Plasminogen activators from the saliva of Desmodus rotundus (common
RT  vampire bat): unique fibrin specificity.";
RL  Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC  -1- FUNCTION: Probably essential to support the feeding habits of this
CC  exclusively haematophagous animal. Probable potent thrombolytic
CC  agent.
CC  -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC  plasminogen to form plasmin.
CC  -1- ENZYME REGULATION: Activity toward plasminogen is stimulated in
CC  the presence of fibrin I.
CC  -1- SUBUNIT: Monomer.
CC  -1- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
CC  and the kringle domain apparently mediates fibrin-induced
CC  stimulation of activity.
CC  -1- SIMILARITY: Belongs to peptidase family S1.
CC  -1- SIMILARITY: Contains 1 EGF-like domain.
CC  -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC  -1- SIMILARITY: Contains 1 kringle domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).

```

```

CC  EMBL; M63988; AAA31593.1; -
DR  EMBL; J05082; AAA31596.1; -
DR  PIR; A34363; A34369.
DR  PIR; J50598; J50598.
DR  HSSP; P98119; IAST.
DR  MEROPS; S01.232; -.
DR  InterPro; IPR009003; Cys_ser_trypsin.
DR  InterPro; IPR006209; EGF-like.
DR  InterPro; IPR000083; Fibrinctn1.
DR  InterPro; IPR006210; IEGF.
DR  InterPro; IPR000001; Kringle_S1.
DR  InterPro; IPR001254; Peptidase_S1A.
DR  Pfam; PF00008; EGF_1.
DR  Pfam; PF00039; fn1; 1.
DR  Pfam; PF00051; kringle; 1.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000395; Kringle; 1.
DR  SMART; SM00058; FN1; 1.
DR  SMART; SM00130; KR; 1.
DR  SMART; SM00020; TYP_Spc; 1.
DR  PROSITE; PS00022; EGF_1; 1.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS50026; EGF_3; 1.
DR  PROSITE; PS01253; FIBRONECTIN_1; 1.
DR  PROSITE; PS00021; KRINGLE_1; 1.
DR  PROSITE; PS50070; KRINGLE_2; 1.
DR  PROSITE; PS50240; TRYPsin_DOM; 1.
DR  PROSITE; PS00134; TRYPsin_HIS; 1.
DR  PROSITE; PS00135; TRYPsin_SER; 1.
KM  Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KM  kringle; EGF-like domain; signal; Multigene family.
FT  SIGNAL 1 36
FT  CHAIN 37 477
FT  DOMAIN 40 82 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.
FT  DOMAIN 83 121 EGF-LIKE.
FT  DOMAIN 128 209 KRINGLE.
FT  DOMAIN 225 477
FT  ACT_SITE 272 272 SERINE PROTEASE.
FT  ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  DISULFID 42 72 BY SIMILARITY.
FT  DISULFID 70 79 BY SIMILARITY.
FT  DISULFID 87 98 BY SIMILARITY.
FT  DISULFID 92 109 BY SIMILARITY.
FT  DISULFID 111 120 BY SIMILARITY.
FT  DISULFID 128 209 BY SIMILARITY.
FT  DISULFID 149 191 BY SIMILARITY.
FT  DISULFID 180 204 BY SIMILARITY.
FT  DISULFID 214 345 BY SIMILARITY.
FT  DISULFID 257 273 BY SIMILARITY.
FT  DISULFID 265 334 BY SIMILARITY.
FT  DISULFID 359 434 BY SIMILARITY.
FT  DISULFID 391 407 BY SIMILARITY.
FT  DISULFID 424 452 BY SIMILARITY.
FT  CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 398 398 N -> K (IN REF. 2).
FT  CONFLICT 403 403 N -> H (IN REF. 2).
FT  CONFLICT 417 417 Y -> R (IN REF. 2).
FT  CONFLICT 435 435 M -> R (IN REF. 2).
SQ  SEQUENCE 477 AA; 53719 MW; 1748655C0E5077C CRC64;

```

```

Query Match      64.5%; Score 40; DB 1; Length 477;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy      2 RPKCV 7
        :|||||
Db      188 KPMCTV 193

```

TPA	TPA MOUSE	STANDARD;	PRT;	559 AA.
AC	pl1214, Q91VW2;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.66) (tPA)			
DE	(t-PA) (t-plasminogen activator).			
GN	PLAT.			
OS	Mus macululus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteleia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=86087303; PubMed=2826484;			
RA	Rickles R.J., Darrow A.L., Strickland S.;			
RT	"Molecular cloning of complementary DNA to mouse tissue plasminogen			
RT	activator mRNA and its expression during F9 teratocarcinoma cell			
RT	differentiation.";			
RL	J. Biol. Chem. 263:1563-1569(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Straubeberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,			
RA	Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,			
RA	Datchenko L., Marina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stepletson M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C.,			
RA	Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Boak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gumarate P.H.,			
RA	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Vallatou D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Faney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting R., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Butterfield A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Rodriguez Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,			
RA	Schmertz A., Schein J.E., Jones S.J.W., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen			
CC	to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By			
CC	controlling plasmin-mediated proteolysis, it plays an important			
CC	role in tissue remodeling and degradation, in cell migration and			
CC	many other physiopathological events.			
CC	-1- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in			
CC	plasminogen to form plasmin.			
CC	-1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide			
CC	bond.			
CC	-1- SUBCELLULAR LOCATION: Secreted; extracellular.			
CC	-1- PFM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER			
CC	PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER			
CC	ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.			
CC	-1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A			
CC	chain. Binding to fibrin enhances its catalytic activity.			
CC	-1- SIMILARITY: Belongs to peptidase family S1.			
CC	-1- SIMILARITY: Contains 1 EGF-like domain.			
CC	-1- SIMILARITY: Contains 1 fibronectin type I domain.			
CC	-1- SIMILARITY: Contains 2 kringle domains.			
CC	-----			
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	-----			

CC or send an email to license@isb.ch).
CC or send an email to license@isb.ch).

DR EMBL; J03520; AAA4070.1; -.
DR EMBL; BC011256; AAH11256.1; -.
DR PIR; A29941; A29941.
DR HSSP; P00750; IASH.
DR MEROPS; S01_232; -.
DR MGD; MG1;97610; Plac.
DR InterPro; IPRO09003; Cys_Ser_trypsin.
DR InterPro; IPRO06209; EGF_like.
DR InterPro; IPRO00083; Fibriactn.
DR InterPro; IPRO06210; IRGF.
DR InterPro; IPRO00001; kingle.
DR InterPro; IPRO01254; peptidase_S1.
DR InterPro; IPRO01314; peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF_1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TYP_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasma; Kringler; EGF-like domain; Repeat; Signal.
KM Plasma; Kringler; EGF-like domain; Repeat; Signal.
FT PROPEP 1 17
FT CHAIN 30 559
FT CHAIN 30 308

FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT_SITE 355 355
FT ACT_SITE 404 404
FT ACT_SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 260 260

TISSUE-TYPE PLASMINOGEN ACTIVATOR.
TISSUE-TYPE PLASMINOGEN ACTIVATOR A
CHAIN.
TISSUE-TYPE PLASMINOGEN ACTIVATOR B
CHAIN.
FIBRONECTIN TYPE-1.
EGF-LIKE.
KRINGLE 1.
KRINGLE 2.
SERINE PROTEASE.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
INTERCHAIN (BY SIMILARITY).
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
G -> A (IN REF. 1).

FT	DISULFID	474	490	BY SIMILARITY.
FT	DISULFID	506	534	BY SIMILARITY.
FT	CARBOHYD	149	149	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	481	481	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	380	380	E -> K (IN REF. 1).
SQ	SEQUENCE	559 AA;	62903 MM;	7DBD3809C1D1C921 CRC64;

Query Match 64.5%; Score 40; DB 1; Length 559;
 Best Local Similarity 83.3%; Pred. No. 14;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPNQYV 7
 Db 184 KPNQYV 189

Search completed: March 18, 2004, 11:17:01
 Job time : 0.612691 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 2.73523 Seconds
(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYVQVQ 10

Scoring table: BLAST62
Gap 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	91.9	154	4	Q96SE8
2	57	91.9	214	6	Q9XT70
3	57	91.9	433	6	Q8MIL0
4	57	91.9	433	6	Q8MHY7
5	56	90.3	157	6	Q9TV48
6	53	85.5	231	11	Q8C6L2
7	46	74.2	1029	5	Q8I252
8	43	69.4	608	16	Q9KOP9
9	43	69.4	653	11	Q8VCS4
10	43	69.4	852	5	Q8IKW0
11	41	66.1	97	10	P82624
12	41	66.1	201	16	Q7U319
13	41	66.1	234	4	Q86TW2
14	41	66.1	263	4	Q00318
15	41	66.1	263	4	Q8NCJ9
16	41	66.1	263	4	Q96FE7

17	41	66.1	517	11	Q8K0D2	Q8K0D2 mus musculus
18	40	64.5	202	13	Q90675	Q90675 gallus gall
19	40	64.5	213	11	Q811Z3	Q811Z3 mus musculus
20	40	64.5	263	13	Q7SXB3	Q7SXB3 brachydactyl
21	40	64.5	264	11	Q811Z2	Q811Z2 mus musculus
22	40	64.5	264	11	Q7TMJ8	Q7TMJ8 mus musculus
23	40	64.5	272	11	Q8BWS2	Q8BWS2 mus musculus
24	40	64.5	291	4	Q7Z7N2	Q7Z7N2 homo sapien
25	40	64.5	304	16	Q9A3Q4	Q9A3Q4 caulobacter
26	40	64.5	411	11	Q8R0L1	Q8R0L1 mus musculus
27	40	64.5	516	4	Q9BU99	Q9BU99 homo sapien
28	40	64.5	519	10	Q8IAC5	Q8IAC5 arabidopsis
29	40	64.5	519	10	Q84M90	Q84M90 arabidopsis
30	40	64.5	560	4	Q14520	Q14520 homo sapien
31	40	64.5	562	4	Q86YK8	Q86YK8 homo sapien
32	40	64.5	564	6	Q8MKB1	Q8MKB1 oryctolagus
33	40	64.5	566	10	Q911L7	Q911L7 arabidopsis
34	40	64.5	567	4	Q13208	Q13208 homo sapien
35	40	64.5	947	13	Q8AXY6	Q8AXY6 gallus gall
36	40	64.5	1109	5	Q869M0	Q869M0 dictyostell
37	39	62.9	60	4	Q9UKJ7	Q9UKJ7 homo sapien
38	39	62.9	113	4	Q9URJ7	Q9URJ7 homo sapien
39	39	62.9	215	13	Q42341	Q42341 gallus gall
40	39	62.9	285	4	Q8TCE2	Q8TCE2 homo sapien
41	39	62.9	280	4	Q02935	Q02935 homo sapien
42	39	62.9	296	4	Q14519	Q14519 homo sapien
43	39	62.9	374	17	Q8Z244	Q8Z244 pyrobaculum
44	39	62.9	421	13	Q8AXX3	Q8AXX3 xenopus lae
45	39	62.9	454	6	Q46506	Q46506 papio hamad

ALIGNMENTS

RESULT 1
ID Q96SE8 PRELIMINARY; PRT; 154 AA.
AC Q96SE8;
DT 01-DEC-2001 (T-REMBLrel. 19, Created)
DT 01-DEC-2001 (T-REMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (T-REMBLrel. 24, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029537; AAK38734.1; -;
DR GO; GO:0016301; F-kinase activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS50070; KRINGLE_2; 1.
KM Glycocytolecin; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17205 MW; A3CF2FCFF505572 CRC64;

Query Match 91.9%; Score 57; DB 4; Length 154;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYQV 9
|||
Db 129 RRPWCYQV 137

RESULT 2

09XT70 PRELIMINARY; PRT; 214 AA.
AC 09XT70; (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA yin J., Idel J. S.;
RT "Partial mRNA of rabbit uPA."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSRP; P00749; IETN.
DR MEROP; S01.231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_Trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
KM Glycocytolecin; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32BD7E CRC64;

Query Match 91.9%; Score 57; DB 6; Length 214;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYQV 9
|||
Db 43 RRPWCYQV 51

RESULT 3

08MIL0 PRELIMINARY; PRT; 433 AA.
AC 08MIL0;
DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.

OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=2215945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
development causes arterial constriction and lumen loss, and
accelerates lesion growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY12285; AAM83187.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_Trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR008295; Pept_S1A_uPA.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR PIRSF; PIRSF001144; Utk_PlasM_act; 1.
KM Glycocytolecin; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match 91.9%; Score 57; DB 6; Length 433;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYQV 9
|||
Db 131 RRPWCYQV 139

RESULT 4

08MHY7 PRELIMINARY; PRT; 433 AA.
AC 08MHY7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;

```

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.,
RT "oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY029517; AK40239.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR01254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR008293; Pept_S1A_upa.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOMA; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Pfam; PF0001144; Uro_kinase; 1.
DR GlycoProtein; Hydroxylase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 4844 MW; 6DD35A371010A6EB CRC64;

Query Match 91.9%; Score 57; DB 6; Length 433;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 131 RRPWCYQV 139

RESULT 5
Q9TVAB PRELIMINARY; PRT; 157 AA.
AC Q9TVAB;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_Taxid=9913;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA MEDLINE=21071388; PubMed=11204721;
RT "Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; -.
DR HSP; P00749; IURK.
DR GO; GO:0016301; F:kinase activity; IEA.

```

```

DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR GlycoProtein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A76BD6C72C1FBFB7 CRC64;

Query Match 90.3%; Score 56; DB 6; Length 157;
Best Local Similarity 88.9%; Pred. No. 0.026;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 97 RRPWCYQV 105

RESULT 6
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasmalogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RC MEDLINE=22354683; PubMed=1246851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RT Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
DR PIR; PT0534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
FT NON_TER 231
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25B8980A682737E2 CRC64;

Query Match 85.5%; Score 53; DB 11; Length 231;
Best Local Similarity 77.8%; Pred. No. 0.13;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 130 RRPWCYQV 138

RESULT 7

```

Q81252
ID 081252 PRELIMINARY; PRT; 1029 AA.
AC 081252;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein with putative LCCL domain.
GN PFA0445W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=36329;
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churche C., Harris B., Harris D., Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooke K., Bucke C.O., Burrows C., Cherevach I., Chillingworth C., Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C., Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J., Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z., Harper D., Hauser H., Hornby T., Holroyd S., Horrocks P., Humphrey S., Jagels K., James K.D., Johnson D., Kerhornou A., Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N., Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L., Oliver K., Ormond D., Price C., Quail M.A., Rabinowitch E., Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M., Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K., Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J., Sultson J.E., Craig A., Newbold C., Barrell B.G., "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.", Nature 419:527-531 (2002).
DR EMBL; AL031745; CAD49049.1; -
DR InterPro; IPR004043; LCCL_dom.
DR Pfam; PF03815; LCCL; 1.
DR SMART; SM00603; LCCL; 1.
DR PROSITE; PS50820; LCCL; 1.
KW Hypothetical protein.
SQ SEQUENCE 1029 AA; 118532 MW; 6512F6CE1B6D3725 CRC64;
Query Match 74.2%; Score 46; DB 5; Length 1029;
Best Local Similarity 77.8%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 RPMCVCVQV 10
Db 73 REMCVCVEVQ 81
RESULT 8
Q9KOP9 PRELIMINARY; PRT; 608 AA.
AC Q9KOP9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PvCA protein.
GN VCI1949.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; *Vibrio*.
NCBI_TaxID=666;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P., McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*

RT cholerae.",
RL Nature 406:477-483 (2000).
DR EMBL; AE004270; AAF95097.1; -
DR PIR; G82137; G82137.
DR TIGR; VCI1949; -
DR InterPro; IPR007817; DIT1_PvCA.
DR Pfam; PF05141; DIT1_PvCA; 1.
KW Complete proteome.
SQ SEQUENCE 608 AA; 70459 MW; BAF8EF3F11C4299 CRC64;
Query Match 69.4%; Score 43; DB 16; Length 608;
Best Local Similarity 62.5%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 2 RPMCVCVQV 9
Db 306 KPMVCYQL 313
RESULT 9
Q8VCS4 PRELIMINARY; PRT; 653 AA.
AC Q8VCS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS *Mus musculus* (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -
DR HSRP; P00761; IAN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEP1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000995; FN_Type_II; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01166; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_2; 1.
DR PROSITE; PS00023; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PSS0070; KRINGLE 2; 1.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
 DR PROSITE; PSS00135; TRYPSIN_SER; 1.
 KM Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
 KM Kringle; Protease; Serine protease.
 SQ SEQUENCE 653 AA; 70553 MW; FE1B090174E6FDD CRC64;

Query Match 69.4%; Score 43; DB 11; Length 653;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPMCIV 7
 DB 343 RPMCIV 348

RESULT 10
 Q8IKW0 PRELIMINARY; PRT; 852 AA.
 AC Q8IKW0;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Hypothetical protein.
 GN PF14_0491.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxId=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=2255705; PubMed=1236864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan W.-S., Nene V., Shallow S.D., Suh B., Peterson J., Angiuoli S.,
 RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrett B.;
 RT "Genome sequence of the human malaria parasite Plasmodium falciparum."
 RL Nature 419:498-511(2002).
 DR EMBL: AB014824; AAA37104.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 852 AA; 98539 MW; 77FD6608825C113C CRC64;

Query Match 69.4%; Score 43; DB 5; Length 852;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RPMCIVQVQ 10
 DB 60 KEMCIVQVQ 68

RESULT 11
 P82624 PRELIMINARY; PRT; 97 AA.
 AC P82624;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE SCRL5 precursor.
 GN SCRL5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxId=3702;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Maltfecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."
 RL Nature 408:816-820(2000).
 RN [2]
 RP IDENTIFICATION, AND TISSUE SPECIFICITY.
 RX PubMed=11437247;
 RA Vanoosthuyse V., Mies G., Dumas C., Cock J.M.;
 RT "Two large Arabidopsis thaliana gene families are homologous to the Brassica gene superfamily that encodes pollen coat proteins and the male component of the self-incompatibility response."
 RL Plant Mol. Biol. 46:17-34(2001).
 CC -1- TISSUE SPECIFICITY: FLOWER BUDS AND STEMS.
 DR EMBL: AC018908; -; NOT_ANNOTATED_CDS.
 KW SIGNAL.
 FT SIGNAL 1 22
 FT CHAIN 23 97 SCRL5.
 SQ SEQUENCE 97 AA; 11119 MW; 66CA0B925B708F54 CRC64;

Query Match 66.1%; Score 41; DB 10; Length 97;
 Best Local Similarity 60.0%; Pred. No. 72;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPMCIVQVQ 10
 DB 35 RRPWCPSKIQ 44

RESULT 12
 Q7U319 PRELIMINARY; PRT; 201 AA.
 AC Q7U319;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Urease accessory protein G.
 GN UREG OR SYN2443.
 OS Synechococcus sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxId=84588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22825697; PubMed=12917641;
 RA Palenik B., Brahamha B., Larimer F.W., Land M., Hausen L.,
 RA Lamerdin J., Regala W., Allen E.E., McCarron J., Paulsen I.,
 RA Dufresne A., Paternsky F., Webb E.A., Waterbury J.;
 RT "The genome of a motile marine Synechococcus."
 RL Nature 424:1037-1042(2003).
 DR EMBL: BX569695; CAB08958.1; -
 KW Complete proteome.
 SQ SEQUENCE 201 AA; 21557 MW; 7096BF2C566CF597 CRC64;

Query Match 66.1%; Score 41; DB 16; Length 201;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RPPWCYVOVQ 10
 |||||:
 Db 174 RPPWCYVNLQ 182

RESULT 13

086YW2 PRELIMINARY; PRT; 234 AA.
 AC 086YW2;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE HGFL(s) protein.
 GN HGFL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chiang H., Chang M.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF528079; AA033762.1; -.
 DR InterPro; IPR000001; Kringie.
 DR Pfam; PF00051; Kringie; 1.
 DR PRINTS; PR00018; KRINGIE.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGIE_1; 1.
 DR PROSITE; PS50070; KRINGIE_2; 1.
 SQ SEQUENCE 234 AA; 25320 MW; C78F64DBD1B8DCDC CRC64;

Query March 66.1%; Score 41; DB 4; Length 234;
 Best Local Similarity 85.7%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPPWCYV 7
 |||||
 Db 78 RPPWCYV 84

RESULT 14
 000318 PRELIMINARY; PRT; 263 AA.
 ID 000318;
 AC 000318;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE WUGSC:DJ515N1.2 protein.
 GN WUGSC:DJ515N1.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Du Z., Scheet P., Harper M.;
 RT "The sequence of H. sapiens PAC clone RP3-515N1."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGIE DOMAIN.
 DR EMBL; AC002073; AAB54054.1; -.
 DR HSSP; P00749; IKDU.
 DR InterPro; IPR000001; Kringie.
 DR Pfam; PF00051; Kringie; 1.
 DR PRINTS; PR00018; KRINGIE.
 DR ProDom; PD000395; Kringie; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGIE_1; 1.
 DR PROSITE; PS50070; KRINGIE_2; 1.

KW Glycoprotein; Kringie.
 SQ SEQUENCE 263 AA; 28248 MW; 197C3EBE8E54A242 CRC64;

Query Match 66.1%; Score 41; DB 4; Length 263;
 Best Local Similarity 85.7%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPPWCYV 7
 |||||
 Db 78 RPPWCYV 84

RESULT 15

086NCJ9 PRELIMINARY; PRT; 263 AA.
 AC 086NCJ9;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein FLJ30207.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Mammary gland;
 RA Isogai T., Ota T., Nishikawa T., Hayaishi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
 RA Yamamoto J., Makamatsu A., Nakamura Y., Kojima S., Nagahara K.,
 RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
 RT "MDO human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGIE DOMAIN.
 DR EMBL; AK074688; BAC11140.1; -.
 DR InterPro; IPR000001; Kringie.
 DR Pfam; PF00051; Kringie; 1.
 DR PRINTS; PR00018; KRINGIE.
 DR ProDom; PD000395; Kringie; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGIE_1; 1.
 DR PROSITE; PS50070; KRINGIE_2; 1.
 KW Hypothetical protein; Glycoprotein; Kringie.
 SQ SEQUENCE 263 AA; 28104 MW; 73A9294D5142CC63 CRC64;

Query March 66.1%; Score 41; DB 4; Length 263;
 Best Local Similarity 85.7%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPPWCYV 7
 |||||
 Db 78 RPPWCYV 84

Search completed: March 18, 2004, 11:19:22
 Job time : 3.73523 secs